

RESULT 12

VNHU

von Willebrand factor precursor - human

C:Species: Homo sapiens (man)

C>Date: 04-Dec-1986 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999

C:Accession: A34480; S02377; A37139; S23676; A25298; A25469; A25366; S23618; S23645; A94

R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Morrall, N.K.; Shelton-Inloes, B.B.; Sor

J. Biol. Chem. 264, 19514-19527, 1989

A:Title: Structure of the gene for human von Willebrand factor.

A:Reference number: A34480; MUID:90062044; PMID:2584182

A:Accession: A34480

A:Molecule type: DNA

A:Residues: 1-2813 <MAN>

A:Cross-references: EMBL:M25864

R:Bonthron, D.; Orkin, S.H.

Eur. J. Biochem. 171, 51-57, 1988

A:Title: The human von Willebrand factor gene. Structure of the 5' region.

A:Reference number: S02377; MUID:88111704; PMID:2828057

A:Accession: S02377

A:Molecule type: DNA

A:Residues: 1-177 <BO2>

A:Cross-references: EMBL:X06828

R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sor

Biochemistry 30, 253-269, 1991

A:Title: Human von Willebrand factor gene and pseudogene: structural analysis and differ

A:Reference number: A37139; MUID:91105089; PMID:1988024

A:Accession: A37139

A:Molecule type: DNA

A:Residues: 990-1947 <MAD>

A:Cross-references: G8:M60675; NID:g340357; PIDN:AAA61295.1; PID:g553810

A>Note: the authors translated the codon CGC for residue 156 as Gln

R:Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagian,

Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987

A:Title: Molecular cloning of the human gene for von Willebrand factor and identification

A:Reference number: S23676; MUID:87260814; PMID:3496594

A:Accession: S23676

A:Molecule type: DNA

A:Residues: 2731-2813 <COL>

A:Cross-references: EMBL:M16945

R:Bonthron, D.; Orr, E.C.; Mitsock, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.

Nucleic Acids Res. 14, 7125-7127, 1986

A:Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.

A:Reference number: A25298; MUID:87016349; PMID:3489923

A:Accession: A25298

A:Molecule type: mRNA

A:Residues: 1-470, 'V', 472-2813 <BON>

A:Cross-references: EMBL:X04385

R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.

EMBO J. 5, 1839-1847, 1986

A:Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protei

A:Reference number: A91044; MUID:87004550; PMID:3019665

A:Accession: A25469

A:Molecule type: mRNA

A:Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>

A:Cross-references: EMBL:X04146

A>Note: this sequence has been revised in reference A91056

R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.

EMBO J. 5, 3074, 1986

A:Reference number: A91056

A:Accession: A25366

A:Molecule type: mRNA

A:Residues: 1021-1030 <VE2>

A>Note: this is a revision to the sequence from reference A91044

R:Shelton-Inloes, B.B.; Broze Jr., G.J.; Miletich, J.P.; Sadler, J.E.

Biochem. Biophys. Res. Commun. 144, 657-665, 1987

A:Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated

A:Reference number: S23618; MUID:87213253; PMID:3495266

A:Accession: S23618

A:Molecule type: mRNA

A:Residues: 1-120 <SH2>

A:Cross-references: EMBL:M17588; NID:g799330; PIDN:AAA65940.1; PID:g340316

A:Accession: S23645

A:Molecule type: protein

A:Residues: 23-56 <SH3>

R:Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Titani, K.; Davie,

Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985

A:Title: Cloning and characterization of two cDNAs coding for human von Willebrand fa

A:Reference number: A94060; MUID:86016708; PMID:2864688

A:Accession: A94060

A:Molecule type: mRNA

A:Residues: 'WA', 739, 'C', 744-769, 'H', 771-788, 'A', 790-803, 'S', 805-873; 1289-1471, 'D', 14

R:Shelton-Inloes, B.B.; Titani, K.; Sadler, J.E.

Biochemistry 25, 3164-3171, 1986

A:Title: cDNA sequences for human von Willebrand factor reveal five types of repeated

A:Reference number: A90504; MUID:86269894; PMID:3488076

A:Accession: A90504

A:Molecule type: mRNA

A:Residues: 781-788, 'A', 790-1424 <SHE>

A>Note: 852-Gln, 857-Asp, and 1381-Thr were also found

R:Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Donlon, T.A.; Bruns, G.A.P.; Latt, S.A.

Science 228, 1401-1406, 1985

A:Title: Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clo

A:Reference number: A44178; MUID:85244588; PMID:3874428

A:Accession: A44178

A:Molecule type: mRNA

A:Residues: 2621-2813 <GIN>

A:Cross-references: EMBL:X03028; NID:g340308; PIDN:AAA61293.1; PID:g340309

R:Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.J.; van Kessel, A.G.;

Nucleic Acids Res. 13, 4699-4717, 1985

A:Title: Construction of cDNA coding for human von Willebrand factor using antibody p

A:Reference number: S07363; MUID:85369603; PMID:3875078

A:Accession: S07363

A:Molecule type: mRNA

A:Residues: 2731-2813 <VE3>

A:Cross-references: EMBL:X02672; NID:g37939; PIDN:CAA26503.1; PID:g37940

R:Lynch, D.C.; Zimmerman, T.S.; Collins, C.J.; Brown, M.; Morfin, M.J.; Ling, E.H.; Li

Cell 41, 49-56, 1985

A:Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by

A:Reference number: S23678; MUID:85201687; PMID:3873280

A:Accession: S23678

A:Molecule type: mRNA

A:Residues: 2731-2813 <LYN>

A:Cross-references: EMBL:X03028

R:Titani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.

Biochemistry 25, 3171-3184, 1986

A:Title: Amino acid sequences of human von Willebrand factor.

A:Reference number: A90505; MUID:86269895; PMID:3524673

A:Accession: A90505

A:Molecule type: protein

A:Residues: 764-788, 'A', 790-1471, 'D', 1473-2813 <TTT>

A>Note: 789-Thr was also found

R:Chopek, M.W.; Girma, J.P.; Fujikawa, K.; Davie, E.W.; Titani, K.

Biochemistry 25, 3146-3155, 1986

A:Title: Human von Willebrand factor: a multivalent protein composed of identical sub

A:Reference number: A23464; MUID:86269892; PMID:3015199

A:Accession: A23464

A:Molecule type: protein

A:Residues: 764-773; 2803-2813 <CHO>

R:Dent, J.A.; Berkowitz, S.D.; Ware, J.; Kasper, C.K.; Ruggeri, Z.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990

A:Title: Identification of a cleavage site directing the immunochemical detection of

A:Reference number: A36013; MUID:90349604; PMID:2385594

A:Accession: A36013

A:Molecule type: protein

A:Residues: 1606-1617 <DEN>

R:Ray, P.J.; Kawai, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B.

Science 232, 995-998, 1986

A:Title: Propolypeptide of von Willebrand factor circulates in blood and is identical

A:Reference number: A60913; MUID:86208144; PMID:3486471

A:Accession: A60913

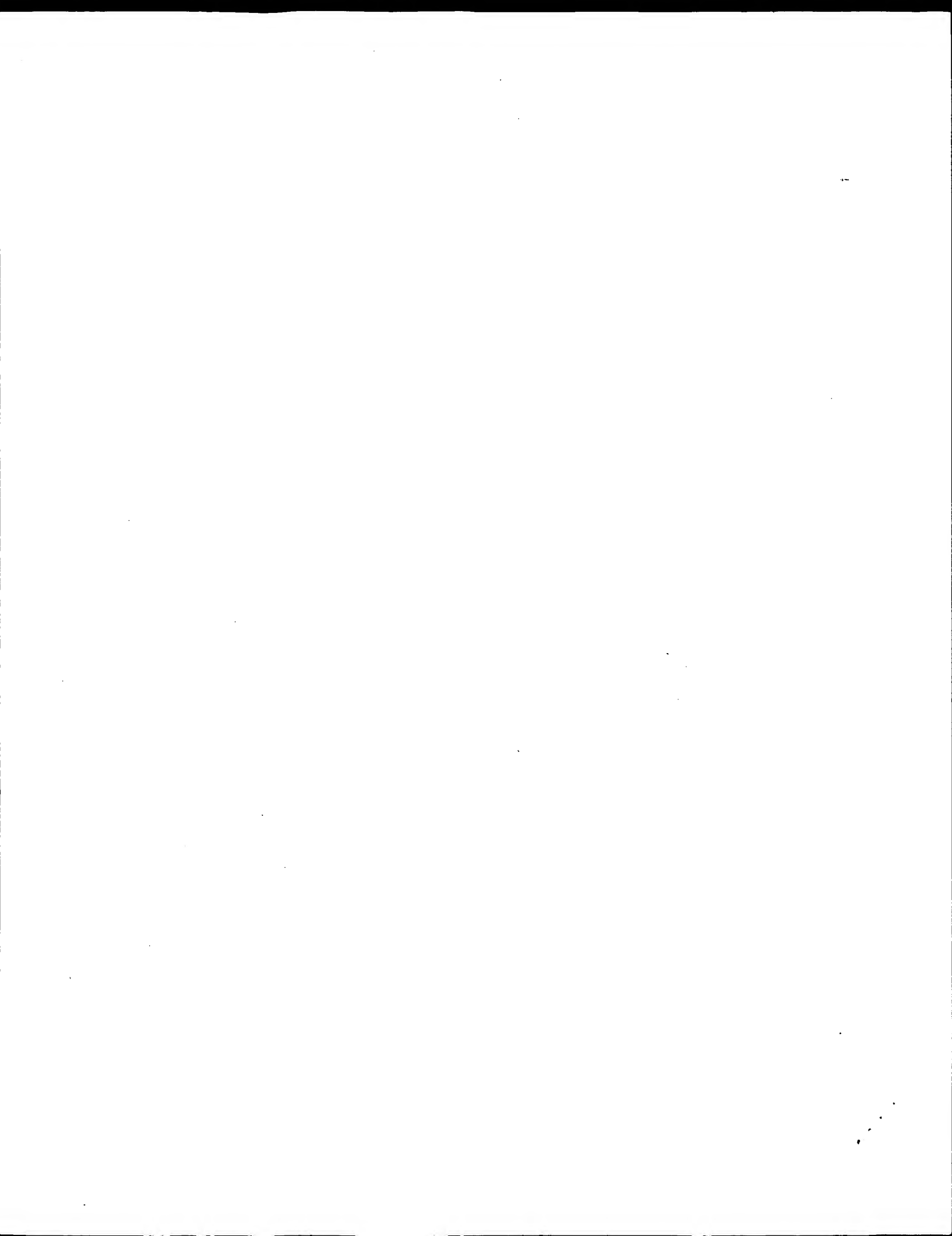
A:Molecule type: protein

A:Residues: 576-590 <FAY>

C:Genetics:

A:Gene: GDB:VWF

A:Cross-references: GDB:119125; OMIM:193400



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:03:22 : Search time 29 Seconds
(without alignments)
95.825 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGLGGRGKCFSEIFSR.....CRGLYLNRKKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	113	28.9	5376	1 ZAN_MOUSE	O88799 mus musculus
2	109	27.9	2700	1 ZAN_HUMAN	Q9y493 homo sapien
3	107.5	27.5	2476	1 ZAN_PIG	Q28983 sus scrofa
4	104.5	26.7	56	1 AMCI_APIPE	P56682 apis mellif
5	103.5	26.5	2813	1 VWF_CANFA	Q28295 canis famli
6	101.5	26.0	937	1 VWF_BOVIN	P80012 bos taurus
7	98.5	25.2	2282	1 ZAN_RABIT	P57999 oryctolagus
8	98	25.1	115	1 A62F_DROME	Q46202 drosophila
9	98	25.1	2813	1 VWF_HUMAN	P04275 homo sapien
10	91.5	23.4	2482	1 VWF_PIG	Q28833 sus scrofa
11	90.5	23.1	3133	1 HMCT_BOMMO	P98092 bombyx mori
12	86	22.0	1700	1 BAR3_CHITE	Q03376 chironomus
13	80.5	20.6	63	1 ICE1_ASCSU	P07851 ascaris suu
14	80.5	20.6	99	1 YOJ2_CAEEL	P34625 caenorhabdi
15	80	20.5	1370	1 IGLR_RAT	P24062 rattus norv
16	80	20.5	1373	1 IGLR_MOUSE	Q60751 mus musculus
17	79.5	20.3	351	1 CRMB_COMPX	Q73559 corpopo viru
18	78.5	20.1	65	1 ICE2_ASCSU	P07852 ascaris suu
19	76.5	19.6	72	1 MT12_MYTED	P80246 mytilus edu
20	76.5	19.6	72	1 MT1B_MYTED	P34625 mytilus edu
21	76.5	19.6	349	1 CRMB_CAMPS	Q8uyv7 camelpox vi
22	76.5	19.6	855	1 ST14_MOUSE	P56677 mus musculus
23	75.5	19.3	349	1 CRMB_VARV	P34015 variola vir
24	75.5	19.3	5179	1 MUC2_HUMAN	Q02817 homo sapien
25	75	19.2	956	1 TSP3_MOUSE	Q05895 mus musculus
26	75	19.2	4289	1 TENX_HUMAN	P22105 homo sapien
27	73	18.7	60	1 MT_PERPL	P52725 perca fluvi
28	72.5	18.5	1416	1 YN81_CAEEL	Q03610 caenorhabdi
29	72	18.4	326	1 VT2_MXAVL	P29825 myxoma viru
30	72	18.4	956	1 TSP3_HUMAN	P49746 homo sapien
31	72	18.4	1367	1 IGLR_HUMAN	P08069 homo sapien
32	71	18.2	1746	1 TENA_PIG	Q29116 sus scrofa
33	70.5	18.0	62	1 ITR1_ASCSU	P19398 ascaris suu

34 70.5 18.0 72 1 MT12_MYTED
35 69 17.6 60 1 MTB_ONCMY
36 69 17.6 60 1 MT_ESOLU
37 69 17.6 77 1 PIF_BOMMO
38 69 17.6 1339 1 ERB3_RAT
39 68.5 17.5 72 1 MT14_MYTED
40 68 17.4 60 1 MT_PLEPL
41 68 17.4 60 1 MT_PSEAM
42 68 17.4 4660 1 LRF2_RAT
43 67.5 17.3 182 1 KRUC_SHEEP
44 67 17.1 62 1 MT4_HUMAN
45 67 17.1 328 1 YRP1_CAEEL

P80247 mytilus edu
P09862 oncorhynchu
P25127 esox lucius
Q10731 bombyx mori
Q62799 rattus norv
P80249 mytilus edu
P07216 pleuronecte
P55945 pseudopleur
P98158 rattus norv
P26372 ovis aries
P47944 homo sapien
Q10043 caenorhabdi

ALIGNMENTS

RESULT 1
ZAN_MOUSE
ID ZAN_MOUSE STANDARD; PRT: 5376 AA.
AC O88799; O08647;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zonadhesin precursor.
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98123114; PubMed=9452463;
RA Gao Z., Garbers D.L.;
RT "Species diversity in the structure of zonadhesin, a sperm-specific
RT membrane protein containing multiple cell adhesion molecule-like
RT domains".
RL J. Biol. Chem. 273:3415-3421(1998).
RN [2]
RP SEQUENCE OF 4864-5376 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=97271566; PubMed=9126492;
RA Gao Z., Harumi T., Garbers D.L.;
RT "Chromosome localization of the mouse zonadhesin gene and the human
RT zonadhesin gene (ZAN)".
RL Genomics 41:119-122(1997).
CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC SIGNALING.
CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC APICAL REGION OF THE SPERM HEAD.
CC -!- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
CC -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
CC ZONA PELLUCIDA.
CC -!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
CC SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
CC -!- DOMAIN: THE VWFD DOMAIN 2 MAY MEDIATE COVALENT
CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
CC -!- SIMILARITY: CONTAINS 25 VWFD DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U97068; AAC26680.1; -.
 DR EMBL; U83190; AAC53125.1; -.
 DR MGD; MGI:106656; ZAN.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003645; FOLN.
 DR InterPro; IPR000998; MAM_domain.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR InterPro; IPR003328; TILa_Cysrich.
 DR InterPro; IPR001007; VWF_C.
 DR InterPro; IPR001846; VWF_D.
 DR Pfam; PF00094; vwd; 4.
 DR Pfam; PF00629; MAM; 3.
 DR Pfam; PF01826; TIL; 25.
 DR Pfam; PF02345; TILa; 25.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00274; FOLN; 11.
 DR SMART; SM00137; MAM; 2.
 DR SMART; SM00214; VWC; 4.
 DR SMART; SM00216; VWD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 18.
 DR PROSITE; PS00740; MAM_1; FALSE_NEG.
 DR PROSITE; PS00060; MAM_2; 3.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 Repeat.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 5376 ZONADHESIN.
 FT DOMAIN 18 5310 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 5311 5337 POTENTIAL.
 FT DOMAIN 5338 5376 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 45 210 MAM 1.
 FT DOMAIN 215 374 MAM 2.
 FT DOMAIN 377 542 MAM 3.
 FT DOMAIN 547 1170 80 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
 (MUCIN-LIKE DOMAIN).
 FT DOMAIN 1171 1280 WVED 1 (PARTIAL).
 FT DOMAIN 1281 1669 WVED 2.
 FT DOMAIN 1670 2056 WVED 3.
 FT DOMAIN 2057 2459 WVED 4.
 FT DOMAIN 2460 2579 WVED 5 (PARTIAL).
 FT DOMAIN 2580 2699 WVED 6 (PARTIAL).
 FT DOMAIN 2700 2819 WVED 7 (PARTIAL).
 FT DOMAIN 2820 2939 WVED 8 (PARTIAL).
 FT DOMAIN 2940 3059 WVED 9 (PARTIAL).
 FT DOMAIN 3060 3179 WVED 10 (PARTIAL).
 FT DOMAIN 3180 3299 WVED 11 (PARTIAL).
 FT DOMAIN 3300 3416 WVED 12 (PARTIAL).
 FT DOMAIN 3417 3536 WVED 13 (PARTIAL).
 FT DOMAIN 3537 3656 WVED 14 (PARTIAL).
 FT DOMAIN 3657 3776 WVED 15 (PARTIAL).
 FT DOMAIN 3777 3892 WVED 16 (PARTIAL).
 FT DOMAIN 3893 4029 WVED 17 (PARTIAL).
 FT DOMAIN 4029 4148 WVED 18 (PARTIAL).
 FT DOMAIN 4149 4263 WVED 19 (PARTIAL).
 FT DOMAIN 4264 4383 WVED 20 (PARTIAL).
 FT DOMAIN 4384 4503 WVED 21 (PARTIAL).
 FT DOMAIN 4504 4623 WVED 22 (PARTIAL).
 FT DOMAIN 4624 4743 WVED 23 (PARTIAL).
 FT DOMAIN 4744 4863 WVED 24 (PARTIAL).
 FT DOMAIN 4864 5261 WVED 25.
 FT DOMAIN 5259 5295 EGF-LIKE.
 FT DISULFID 5263 5274 BY SIMILARITY.
 FT DISULFID 5268 5283 BY SIMILARITY.
 FT DISULFID 5285 5299 BY SIMILARITY.
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1216 1216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1239 1239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1314 1314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1814 1814 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1908 1908 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1933 1933 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2028 2028 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 2111 2111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2142 2142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2332 2332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2533 2533 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2575 2575 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2692 2692 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2812 2812 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3052 3052 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3065 3065 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3144 3144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3172 3172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3288 3288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3292 3292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3782 3782 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4005 4005 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4136 4136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4243 4243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4254 4254 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4335 4335 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4376 4376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4586 4586 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5136 5136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5252 5252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 5376 AA; 579908 MW; 0E44DB77DF2A2620 CRC64;
 Query Match 28.9%; Score 113; DB 1; Length 5376;
 Best Local Similarity 37.5%; Pred. No. 0.0034;
 Matches 24; Conservative 8; Mismatches 22; Indels 10; Gaps 3;
 QY 11 KCPSEIFSRCDGRCQFCFNVVPRPLC-----IKICAPCGVGLGVLRNKKVCYVPR 63
 Db 4743 KCPANSLYTHCLPTCLPSCSN--PDGRCEGTSHKAPSTCREGCVQPGYLLN-KDTCVHK 4799
 QY 64 SKCG 67
 Db 4800 NQCG 4803
 RESULT 2
 ID ZAN_HUMAN STANDARD; PRT; 2700 AA.
 AC Q9Y493; O00218;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zonadhesin (Fragment).
 GS ZAN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-2379 FROM N.A.
 RX MEDLINE=99018118; PubMed=9799793;
 RA Glockner G., Scherer S., Schattevoy R., Boright A., Weber J.,
 RA Tsui L.C., Rosenthal A.;
 RT "Large-scale sequencing of two regions in human chromosome 7q22:
 RT analysis of 650 kb of genomic sequence around the EPO and CUTLI loci
 RT reveals 17 genes.";
 RL Genome Res. 8:1060-1073(1998).
 RN [2]
 RP SEQUENCE OF 2338-2700 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=97271566; PubMed=9126492;
 RA Gao Z., Harumi T., Garbers D.L.;
 RT "Chromosome localization of the mouse zonadhesin gene and the human
 RL zonadhesin gene (ZAN).";
 RL Genomics 41:119-122(1997).
 CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE

CC CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).

CC CC -!- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.

CC CC -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE

CC CC ZONA PELLUCIDA.

CC CC -!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,

CC CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF

CC CC SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.

CC CC -!- DOMAIN: THE VWFD DOMAINS 2 AND 3 MAY MEDIATE COVALENT

CC CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).

CC CC -!- SIMILARITY: CONTAINS 3.5 MAM DOMAINS.

CC CC -!- SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC CC the European Bioinformatics Institute. There are no restrictions on its

CC CC use by non-profit institutions as long as its content is in no way

CC CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC -----

CC CC EMBL; AF053356; AAC78790.1; -.

CC CC EMBL; U83191; AAC51208.1; -.

CC CC Genew; HGNC:12857; ZAN.

CC CC MIM: 602372; -.

CC CC InterPro; IPR000561; EGF-like.

CC CC InterPro; IPR000998; MAM_domain.

CC CC InterPro; IPR002919; TIL_Cysrich.

CC CC InterPro; IPR003328; TILA_Cysrich.

CC CC InterPro; IPR001846; VWF_D.

CC CC Pfam; PF00094; vwd; 4.

CC CC Pfam; PF00629; MAM; 4.

CC CC Pfam; PF01826; TIL; 5.

CC CC Pfam; PF02345; TILA; 4.

CC CC SMART; SM00216; VWD; 1.

CC CC PROSITE; PS01186; EGF-2; 3.

CC CC PROSITE; PS00740; MAM_1; 1.

CC CC PROSITE; PS00600; MAM_2; 4.

CC CC Glycoprotein; Transmembrane; Cell adhesion; Repeat.

CC CC NON_TER 1 1

CC CC FT DOMAIN <1 109 MAM 1.

CC CC FT DOMAIN 112 136 MAM 2 (PARTIAL).

CC CC FT DOMAIN 161 326 MAM 3.

CC CC FT DOMAIN 322 446 MAM 4.

CC CC FT DOMAIN 483 951 66 X HEPTAPEPTIDE REPEATS (APPROXIMATE)

CC CC (MUCIN-LIKE DOMAIN).

CC CC FT DOMAIN 953 1065 VWFD 1 (PARTIAL).

CC CC FT DOMAIN 1066 1454 VWFD 2.

CC CC FT DOMAIN 1455 1861 VWFD 3.

CC CC FT DOMAIN 1862 2292 VWFD 4.

CC CC FT DOMAIN 2293 2684 VWFD 5.

CC CC FT DOMAIN ? ? EGF-LIKE.

CC CC FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 403 403 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 1023 1023 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 1099 1099 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 1618 1618 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 1737 1737 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 1832 1832 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 1878 1878 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 2136 2136 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CARBOHYD 2505 2505 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CC FT CONFLICT 2374 2379 NQKMA -> RAGPGP (IN REF. 1).

CC CC FT NON_TER 2700 2700

CC CC SQ SEQUENCE 2700 AA; 293013 MW; 80E60CC0B12277B1 CRC64;

Query Match 27.9%; Score 109; DB 1; Length 2700;

Best Local Similarity 35.3%; Pred. No. 0.00052;

Matches 24; Conservative 11; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPSENFISRC-----DGRQFCFNNVVKPLCKIKCAPGCVGLRNKKV 59

DB 2173 ECPAYSSYTNCLPSCSPSCWDLGRCE--GAKVP-----SACAGCICQPGVLSLSEDK- 2223

QY 60 CVPRSKCG 67

DB 2224 CVPRSKCG 2231

RESULT 3

ZAN_PIG

ID ZAN_PIG STANDARD; PRT; 2476 AA.

AC Q28983;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Zonadhesin precursor.

GN ZAN.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 823-830; 859-872; 883-890;

RP 920-925; 960-967; 1235-1244; 1349-1354; 1518-1532; 1624-1656;

RP 1658-1667; 1777-1795 AND 1914-1921.

RC STRAIN=Weishan; TISSUE=Testis;

RX MEDLINE=96064658; PubMed=7592795;

RA Hardy D.M., Garbers D.L.;

RT "A sperm membrane protein that binds in a species-specific manner to

RT the egg extracellular matrix is homologous to von Willebrand

RT factor";

RL J. Biol. Chem. 270:26025-26028(1995).

CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA

CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR

CC SIGNALING.

CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE

CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).

CC -!- TISSUE SPECIFICITY: IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS.

CC NOT IN LUNG, LIVER, HEART, SPLEEN, BRAIN, KIDNEY, EPIDIDYMIS.

CC -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE

CC ZONA PELLUCIDA.

CC -!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,

CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF

CC SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.

CC -!- DOMAIN: THE VWFD DOMAINS 2 AND 3 MAY MEDIATE COVALENT

CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).

CC -!- PTM: THE MAM DOMAINS AND THE MUCIN-LIKE DOMAINS ARE MISSING

CC FROM THE ZONADHESIN THAT BINDS TO THE EGG EXTRACELLULAR MATRIX.

CC PROCESSING MIGHT OCCUR DURING SPERM MATURATION AND/OR

CC CAPACITATION.

CC -!- SIMILARITY: CONTAINS 2 MAM DOMAINS.

CC -!- SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U40024; AAC48486.1; -.

HSSP; P56682; ICCV.

InterPro; IPR000561; EGF-like.

InterPro; IPR000998; MAM_domain.

InterPro; IPR002919; TIL_Cysrich..

InterPro; IPR003328; TILA_Cysrich..

InterPro; IPR001007; VWF_C.

InterPro; IPR001846; VWF_D.

Pfam; PF00094; vwd; 4.

Pfam; PF00629; MAM; 2.

Pfam; PF01826; TIL; 5.

Pfam; PF02345; TILA; 5.

SMART; SM00181; EGF; 1.

DR SMART; SM00137; MAM; 1.
 DR SMART; SM00214; VWC; 2.
 DR SMART; SM00216; VWD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS00060; MAM_2; 2.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 KW Repeat.
 FT CHAIN 1 29
 FT CHAIN 30 2476
 FT DOMAIN 30 2418
 FT TRANSDOM 2419 2439
 FT DOMAIN 2440 2476
 FT DOMAIN 31 144
 FT DOMAIN 147 312
 FT DOMAIN 319 687
 FT DOMAIN 688 799
 FT DOMAIN 800 1184
 FT DOMAIN 1185 1573
 FT DOMAIN 1574 1968
 FT DOMAIN 1969 2370
 FT DOMAIN 2366 2402
 FT DISULFID 2370 2381
 FT DISULFID 2375 2390
 FT DISULFID 2392 2401
 FT CARBOHYD 109 109
 FT CARBOHYD 269 269
 FT CARBOHYD 735 735
 FT CARBOHYD 758 758
 FT CARBOHYD 833 833
 FT CARBOHYD 1154 1154
 FT CARBOHYD 1329 1329
 FT CARBOHYD 1448 1448
 FT CARBOHYD 1544 1544
 FT CARBOHYD 1596 1596
 FT CARBOHYD 1654 1654
 FT CARBOHYD 1843 1843
 FT CARBOHYD 1965 1965
 FT CARBOHYD 2122 2122
 FT CARBOHYD 2165 2165
 FT CARBOHYD 2178 2178
 FT CARBOHYD 2329 2329
 FT CARBOHYD 2359 2359
 FT CONFLICT 823 823
 FT CONFLICT 923 923
 FT CONFLICT 965 965
 FT CONFLICT 1241 1241
 SQ SEQUENCE 2476 AA; 270364 MW; 27.58; Score 107.5; DB 1; Length 2476;
 Query Match 27.58; Score 107.5; DB 1; Length 2476;
 Best Local Similarity 33.3%; Pred. No. 0.00071;
 Matches 22; Conservative 9; Mismatches 18; Indels 17; Gaps 3;
 QY 11 KCPNSIEFSRCDGRC-----ORFCNVVVPKLCIKCAPGCVCRGLGYLRNKKVKCV 61
 DB 1455 KCPGSSYSTCANPCPATCLSLNPNYCPSYLP-----CAGCEQCKGHILSGTS-CV 1506
 QY 62 PRSKCG 67
 DB 1507 PLUSQCG 1512
 RESULT 4
 ID AMCI_APIME STANDARD; PRT; 56 AA.
 AC P56682;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Chymotrypsin inhibitor (AMCI).
 OS Apis mellifera (Honeybee).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Aculeata; Apoidea; Apidae; Apis.
 OX NCBI_TaxID=7460;
 RN
 RP SEQUENCE, AND STRUCTURE BY NMR.
 RC TISSUE=Hemolymph;
 RX MEDLINE=99339935; PubMed=10411628;
 RA Bania J., Stachowiak D., Polanowski A.;
 RT "Primary structure and properties of the cathepsin G/chymotrypsin
 RT inhibitor from the larval hemolymph of Apis mellifera.";
 RL Eur. J. Biochem. 262:680-687(1999).
 CC -!- FUNCTION: CHYMOTRYPSIN AND CATHEPSIN G INHIBITOR.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 DR PDB; ICCV; 12-MAR-99.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR Pfam; PF01826; TIL; 1.
 KW Serine protease inhibitor; 3D-structure.
 FT DISULFID 3 36
 FT DISULFID 12 32
 FT DISULFID 16 28
 FT DISULFID 20 56
 FT DISULFID 38 50
 SQ SEQUENCE 56 AA; 5973 MW; 092B2815AE6B2B7F CRC64;
 Query Match 26.7%; Score 104.5; DB 1; Length 56;
 Best Local Similarity 36.8%; Pred. No. 7.6e-05;
 Matches 21; Conservative 7; Mismatches 26; Indels 3; Gaps 2;
 QY 11 KCPNSIEFSRCDGRCFCQFCQPNVVPKP-LCIKICAPGCVCRGLGYLRNKKVKCVPRSKC 66
 DB 2 ECGPNEVFNTCGSACAPTCAQ--PKTRICTMQCRICGCGFLRNGEGACVLPENC 56
 RESULT 5
 ID VWF_CANFA STANDARD; PRT; 2813 AA.
 AC Q28295; Q28311; Q9TS14;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 18-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Von Willebrand factor precursor (VWF).
 GN F8VWF OR VWF.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN
 RP SEQUENCE FROM N.A.
 RA Stoy S.J., Shibuya H., Nonneman D.J., Holzhauser J., Mohammed I.H.,
 RA Johnson G.S.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Montgomery R.R., Fahs S., Montgomery M.W.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Venta P.J., Li J., Yuzbasiyan-Gurkan V., Brewer G.J., Schall W.D.;
 RT "Complete sequence of the structural gene for canine von Willebrand
 RT factor and identification of a mutation causing Scottish terrier von
 RT Willebrand's disease.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1234-1669 FROM N.A.
 RC TISSUE=Blood;
 RA Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
 RT "The canine von Willebrand factor gene: sequence and expression of
 RT a region encoding the glycoprotein Ib/IX binding domain.";
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS. IT
 CC PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
 CC NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF

DR	QY	DB	RESULT	8	AG2F_IDROME	AD	A62F_DROME	STANDARD	PRT	115	AA.	Query Match	Best Local Similarity	25.2%	Score 98.5	DB 1	Length 2282	Matches	22	Conservative	9	Mismatches	25	Indels	3	Gaps	
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P
DR	QY	DB	11	KCP	SP	NEI	SR	DG	RQ	R	CP	NW	-	V	P	R	P	L	C	K	I	-	-	-	-	-	-
DR	QY	DB	1266	KCP	PN	SS	Y	SG	SP	G	CG	T	P	G	T	SL	N	H	P	K	D	C	P	T	I	L	P

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: U85763; AAB96387.1; -
 EMBL: AE003475; AAR47683.1; -
 EMBL: AY010608; AAG35367.1; -
 EMBL: AY010609; AAG35368.1; -
 EMBL: AY010610; AAG35369.1; -
 EMBL: AY010611; AAG35370.1; -
 EMBL: AY010612; AAG35371.1; -
 EMBL: AY010613; AAG35372.1; -
 EMBL: AY010614; AAG35373.1; -
 EMBL: AY010615; AAG35374.1; -
 EMBL: AY010616; AAG35375.1; -
 EMBL: AY010617; AAG35376.1; -
 FlyBase: FBgn020509; Acp62F.
 InterPro: IPR002919; rTL_Cysrich.
 Pfam: PF01826; TIL; 1.
 Behavior: Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 115 ACCESSORY GLAND PROTEIN ACP62F.
 SQ SEQUENCE 115 AA; 12570 MW; 4326AA6FC32291D CRC64;

Query Match 25.18; Score 98; DB 1; Length 115;
 Best Local Similarity 34.48; Pred. No. 0.00065;
 Matches 22; Conservativity 9; Mismatches 31; Indels 2; Gaps 2;

QY 4 GGLGRGKCPNIEFSGDGRGRCPCFNVVPEKPLKIKICAPGCVCRGLGYLRNKK-KVCVP 62
 DB 26 GQQGPKVDTANGTQTECPVACPCEYSGNGP-CVKMGAPCVCKPGYVNERIPACVL 84

QY 63 RSKC 66

DB 85 RSDC 88

RESULT 9

VWF_HUMAN STANDARD; PRT; 2813 AA.
 AC P04275;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Von Willebrand factor precursor (vWF).
 GN F8VWF OR VWF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90062044; PubMed=2584182;
 RA Mancuso D.J., Tuley E.A., Westfield L.A., Worrall N.K.,
 RA Shelton-Inloes B.B., Sorace J.M., Alevis Y.G., Sadler J.E.;
 RT "Structure of the gene for human von Willebrand factor.";
 RL J. Biol. Chem. 264:19514-19527(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87016349; PubMed=3489923;
 RA Bonthron D., Orr E.C., Mitsock L.M., Ginsburg D., Handin R.I.,
 RA Orkin S.H.;
 RT "Nucleotide sequence of pre-pro-von Willebrand factor cDNA.";
 RL Nucleic Acids Res. 14:7125-7128(1986).
 RN [3]
 RP SEQUENCE OF 1-120 FROM N.A., AND SEQUENCE OF 23-56.
 RC TISSUE-Umbilical vein endothelial cells;
 RX MEDLINE=87213253; PubMed=3495266;
 RA Shelton-Inloes B.B., Broze G.J. Jr., Miletich J.P., Sadler J.E.;

RT "Evolution of human von Willebrand factor: cDNA sequence
 RT polymorphisms, repeated domains, and relationship to von Willebrand
 RT antigen II.";
 RL Biochem. Biophys. Res. Commun. 144:657-665(1987).
 RN [4]
 RP SEQUENCE OF 1-1400 FROM N.A.
 RX MEDLINE=87004550; PubMed=3019665;
 RA Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
 RT "Full-length von Willebrand factor (vWF) cDNA encodes a highly
 RT repetitive protein considerably larger than the mature vWF subunit.";
 RL EMBO J. 5:1839-1847(1986).
 RN [5]
 RP ERRATUM.
 RA Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
 RL EMBO J. 5:3074-3074(1986).
 RN [6]
 RP SEQUENCE OF 764-2813.
 RX MEDLINE=86269895; PubMed=3524673;
 RA Titani K., Kumar S., Takio K., Ericsson L.H., Wade R.D., Ashida K.,
 RA Walsh K.A., Choek M.W., Sadler J.E., Fujikawa K.;
 RT "Amino acid sequence of human von Willebrand factor.";
 RL Biochemistry 25:3171-3184(1986).
 RN [7]
 RP SEQUENCE OF 781-1424 FROM N.A.
 RX MEDLINE=86269894; PubMed=3488076;
 RA Shelton-Inloes B.B., Titani K., Sadler J.E.;
 RT "cDNA sequences for human von Willebrand factor reveal five types of
 RT repeated domains and five possible protein sequence polymorphisms.";
 RL Biochemistry 25:3164-3171(1986).
 RN [8]
 RP SEQUENCE OF 764-873 AND 1289-2813 FROM N.A.
 RX MEDLINE=86016708; PubMed=2864688;
 RA Sadler J.E., Shelton-Inloes B.B., Sorace J.M., Harlan J.M.,
 RA Titani K., Davie E.W.;
 RT "Cloning and characterization of two cDNAs coding for human von
 RT Willebrand factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6394-6398(1985).
 RN [9]
 RP SEQUENCE OF 990-1947 FROM N.A.
 RX MEDLINE=91105089; PubMed=1988024;
 RA Mancuso D.J., Tuley E.A., Westfield L.A., Lester-Mancuso T.L.,
 RA Le Beau M.M., Sorace J.M., Sadler J.E.;
 RT "Human von Willebrand factor gene and pseudogene: structural analysis
 RT and differentiation by polymerase chain reaction.";
 RL Biochemistry 30:253-269(1991).
 RN [10]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=85269603; PubMed=3875078;
 RA Verweij C.L., de Vries C.J.M., Distel B., van Zonneveld A.-J.,
 RA Geurts van Kessel A., van Mourik J.A., Pannekoek H.;
 RT "Construction of cDNA coding for human von Willebrand factor using
 RT antibody probes for colony-screening and mapping of the chromosomal
 RT gene.";
 RL Nucleic Acids Res. 13:4699-4717(1985).
 RN [11]
 RP SEQUENCE OF 1-177 FROM N.A.
 RX MEDLINE=88111704; PubMed=2828057;
 RA Bonthron D., Orkin S.H.;
 RT "The human von Willebrand factor gene. Structure of the 5' region.";
 RL Eur. J. Biochem. 171:51-57(1988).
 RN [12]
 RP SEQUENCE OF 2621-2813 FROM N.A.
 RX MEDLINE=85244588; PubMed=3874428;
 RA Ginsburg D., Handin R.I., Bonthron D.T., Donlon T.A., Bruns G.A.P.,
 RA Latt S.A., Orkin S.H.;
 RT "Human von Willebrand factor (vWF): isolation of complementary DNA
 RT (cDNA) clones and chromosomal localization.";
 RL Science 228:1401-1406(1985).
 RN [13]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=85201687; PubMed=3873280;
 RA Lynch D.C., Zimmerman T.S., Collins C.J., Brown M., Morin M.J.,
 RA Ling E.H., Livingston D.M.;


```

CC CC      -!- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
DR DR      EMBL: AF052036; AAC06229.1; -
DR DR      EMBL: S64541; AAB27829.2; -
DR DR      HSSP: P04275; IATZ.
DR DR      InterPro: IPR000359; Cys_knot.
DR DR      InterPro: IPR000561; EGF-like.
DR DR      InterPro: IPR001928; Endothln_tox.
DR DR      InterPro: IPR000864; Potato_inhibit.
DR DR      InterPro: IPR002919; TIL_Cysrich.
DR DR      InterPro: IPR002035; VWF_A.
DR DR      InterPro: IPR001007; VWF_C.
DR DR      InterPro: IPR001846; VWF_D.
DR DR      Pfam: PF00007; Cys_knot; 1.
DR DR      Pfam: PF00092; waa; 3.
DR DR      Pfam: PF00093; wvc; 3.
DR DR      Pfam: PF00094; vwd; 3.
DR DR      Pfam: PF01826; TIL; 3.
DR DR      PRINTS: PR00365; ENDOTHELIN.
DR DR      PRINTS: PR00292; POTATOINHBT.
DR DR      PRINTS: PR00453; VWFADOMAIN.
DR DR      SMART: SM00041; CT; 1.
DR DR      SMART: SM00181; EGF; 1.
DR DR      SMART: SM00327; VWA; 3.
DR DR      SMART: SM00214; VWC; 5.
DR DR      SMART: SM00216; VWD; 3.
DR DR      PROSITE: PS02334; WVEA; 3.
DR DR      PROSITE: PS01208; WVEC; 3.
DR DR      PROSITE: PS01185; CTCK_1; 1.
DR DR      PROSITE: PS01225; CTCK_2; 1.
KW KW      Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW KW      Plasma; Hemostasis; Repeat; Cell adhesion.
FT FT      NON_TER 1 1
FT FT      PROPEP <1 437 BY SIMILARITY.
FT FT      CHAIN 438 2482 VON WILLEBRAND FACTOR.
FT FT      DOMAIN 62 215 WVED 2.
FT FT      DOMAIN 438 461 AMINO-TERMINAL.
FT FT      DOMAIN 462 507 E1.
FT FT      DOMAIN 500 527 CX.
FT FT      DOMAIN 541 687 WVED 3.
FT FT      DOMAIN 947 1127 WVEA 1.
FT FT      DOMAIN 1167 1334 WVEA 2.
FT FT      DOMAIN 1360 1540 WVEA 3.
FT FT      DOMAIN 1619 1771 WVEA 4.
FT FT      DOMAIN 1885 1930 E2.
FT FT      DOMAIN 1924 1997 WVEC 1.
FT FT      DOMAIN 2098 2164 WVEC 2.
FT FT      DOMAIN 2249 2319 WVEC 3.
FT FT      DOMAIN 2393 2481 CTCK.
FT FT      SITE 2176 2178 CELL ATTACHMENT SITE (POTENTIAL).
FT FT      DISULFID 441 482 BY SIMILARITY.
FT FT      DISULFID 450 478 BY SIMILARITY.
FT FT      DISULFID 484 495 BY SIMILARITY.
FT FT      DISULFID 541 670 BY SIMILARITY.
FT FT      DISULFID 563 705 BY SIMILARITY.
FT FT      DISULFID 572 667 BY SIMILARITY.
FT FT      DISULFID 588 595 BY SIMILARITY.
FT FT      DISULFID 734 758 BY SIMILARITY.
FT FT      DISULFID 745 785 BY SIMILARITY.
FT FT      DISULFID 763 765 BY SIMILARITY.
FT FT      DISULFID 827 839 BY SIMILARITY.
FT FT      DISULFID 823 843 BY SIMILARITY.
FT FT      DISULFID 800 804 BY SIMILARITY.
FT FT      DISULFID 870 873 BY SIMILARITY.
FT FT      DISULFID 908 911 BY SIMILARITY.

```

```

FT FT      DISULFID 942 1128 BY SIMILARITY.
FT FT      DISULFID 1338 1339 BY SIMILARITY.
FT FT      DISULFID 1355 1541 BY SIMILARITY.
FT FT      DISULFID 1548 1573 BY SIMILARITY.
FT FT      DISULFID 1568 1609 OR 1611 (BY SIMILARITY).
FT FT      DISULFID 1641 1792 BY SIMILARITY.
FT FT      DISULFID 1619 1754 BY SIMILARITY.
FT FT      DISULFID 1596 1757 BY SIMILARITY.
FT FT      DISULFID 1662 1670 BY SIMILARITY.
FT FT      DISULFID 2393 2443 BY SIMILARITY.
FT FT      DISULFID 2408 2457 BY SIMILARITY.
FT FT      DISULFID 2419 2473 BY SIMILARITY.
FT FT      DISULFID 2423 2475 BY SIMILARITY.
FT FT      DISULFID ? 2480 BY SIMILARITY.
FT FT      CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 905 905 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 1184 1184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 1243 1243 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 1640 1640 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 1892 1892 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 1959 1959 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 2026 2026 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 2069 2069 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 2215 2215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 2254 2254 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      CARBOHYD 2459 2459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT      SEQUENCE 2482 AA; 272394 MW; D499B7DDFBCEADD CRC64;
SQ SEQUENCE 2482 AA; 272394 MW; D499B7DDFBCEADD CRC64;

Query Match 23.4%; Score 91.5; DB 1; Length 2482;
Best Local Similarity 28.8%; Pred. No. 0.035;
Matches 19; Conservative 11; Mismatches 31; Indels 5; Gaps 2;

QY 2 GFGGLGGRGKCFNSNEIFSCDRCQRFQPNV-VPKPLKIKICAPGCVRLGYLRNKKVC 60
Db 320 GFCAL----SCPPGQVYLCQGTFCNLTCRSLSPDECAEDCLGFCFPPGLYLDGSGDC 375
QY 61 VPRSKC 66
Db 376 VPKAQC 381

RESULT 11
HMCT_BOMMO
ID HMCT_BOMMO STANDARD; PRT; 3133 AA.
AC P98092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemocytin precursor (Humoral lectin).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Fuyou X Tokai; TISSUE-Hemocyte;
RX MEDLINE=95178544; PubMed=7873598;
RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
RA Matsubara F., Tanai K., Kadono-Okada K., Kato Y., Mori H.;
RT Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
RT "Cloning and expression of the gene of hemocytin, an insect humoral
RT lectin which is homologous with the mammalian von Willebrand
RT factor.";
RL Blochim. Biophys. Acta 1260:245-258(1995).
RN [2]
RP SEQUENCE OF 2221-3133 FROM N.A.
RA Kotani E., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
RA Matsubara F., Yamakawa M.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR
CC ENCAPSULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL
CC METAMORPHOSIS.

```



```

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52263; CAA36506.1; .
DR PIR: S08167; S08167.
DR HSSP: P15358; 1SKZ.
DR InterPro: IPR004153; CXCXC_repeat.
DR Pfam: PF03128; CXCXC; 71.
KW Repeat; Signal.
FT CHAIN 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202828521B0815 CRC64;
Query Match 22.08; Score 86; DB 1; Length 1700;
Best Local Similarity 29.1%; Pred. No. 0.099;
Matches 23; Conservative 9; Mismatches 21; Indels 26; Gaps 6;
QY 12 CPSNEIFSRDGRQRCPCPNVVKP-----LCIKICAPGCV---CRILGYLRNK 56
DB 1234 CPGNOI--CDTCRCVCPKMKRPAIDCKTKWNDEMCCVCCKPSPGEGCK-GVMKN 1290
QY 57 KKVC---VPRSK-----CG 67
DB 1291 ANTSCCECPADKAKPASC 1309
RESULT 13
ID ICEL_ASCSU STANDARD; PRT; 63 AA.
AC P07851;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chymotrypsin/elastase isoform 1 (C/E-1 inhibitor).
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoideae).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=84255715; PubMed=6564898;
RA Babin D.R., Peanasky R.J., Coos S.M.;
RT "The isoforms of chymotrypsin/elastase from Ascaris
RL lumbricoideae: the primary structure.";
RL Arch. Biochem. Biophys. 232:143-161(1984).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH ELASTASE.
RX MEDLINE=95006335; PubMed=7922044;
RA Huang K., Strynadka N.C., Bernard V.D., Peanasky R.J., James M.N.;
RT "The molecular structure of the complex of Ascaris
RT chymotrypsin/elastase inhibitor with porcine elastase.";
RL Structure 2:679-689(1994).
CC -!- FUNCTION: DEFEND THE ORGANISM AGAINST THE HOST'S PROTEINASES.
CC -!- SIMILARITY: BELONGS TO THE ASCARIS TRYPSIN INHIBITOR FAMILY.
DR PIR: S07127; S07127.
DR PDB: IEAL; 05-APR-99.
DR InterPro: IPR002919; TIL_Cysrich.
DR Pfam: PF01826; TIL; 1.
KW Serine protease inhibitor; 3D-structure.
FT DISULFID 5 38
FT DISULFID 14 33
FT DISULFID 17 29
FT DISULFID 21 60
FT DISULFID 40 54

```

```

FT ACT_SITE 31 32 REACTIVE BOND.
SQ SEQUENCE 63 AA; 6862 MW; 5DC10DE75B375F16 CRC64;
Query Match 20.6%; Score 80.5; DB 1; Length 63;
Best Local Similarity 30.3%; Pred. No. 0.029;
Matches 20; Conservative 10; Mismatches 23; Indels 13; Gaps 4;
QY 8 GRCKCPSNEIFSRDGRQRCPCPNVVKPCLIKICAPGCVCRILGYLRNKKKVC 60
DB 1 GQSCSPNEVWTECTCEMKCGPDENTPCPLMCRRPSC--ECSPG-----RGMRRNTDNGKC 54
QY 61 VPRSKC 66
DB 55 IPASQC 60
RESULT 14
ID YOJ2_CAEEL STANDARD; PRT; 99 AA.
AC F34625;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 10.6 kDa protein ZK353.2 in chromosome III.
GN ZK353.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins J., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [1]
RP SEQUENCE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L15313; AAA28199.1; .
DR WormPep: ZK353.2; CE00386.
KW Hypothetical protein.
SQ SEQUENCE 99 AA; 10561 MW; 862C659838E47E5F CRC64;
Query Match 20.6%; Score 80.5; DB 1; Length 99;
Best Local Similarity 28.1%; Pred. No. 0.041;
Matches 18; Conservative 10; Mismatches 17; Indels 19; Gaps 2;
QY 1 GGFGLGGKGRKPSNEIFSRDGRQRCPCPN-----VPRKPLCIKICA 43
DB 31 GGYGGYGGYGGCGGADNVFVR--WRCCDYSPECCIQLETFWVFLVFIIGFFVCLACL 88
QY 44 PGCV 47
DB 89 AGCV 92

```




QY 62 -PRSK 65
 DB 112 CPRTK 116

RESULT 2

Q9U1T6 PRELIMINARY; PRT; 561 AA.
 AC Q9U1T6;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Y69H2.3a protein.
 DE Y69H2.3B.
 GN Caenorhabditis elegans.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.A.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RT Science 282:2012-2018(1998).
 DR EMBL; 298877; CAB54472.1; -.
 DR HSSP; P56682; ICCV.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003645; FOLN.
 DR Pfam; PF01826; TIL; 3.
 DR SMART; SM00274; FOLN; 2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
 SQ SEQUENCE 561 AA; 60751 MW; 4693023355972B2 CRC64;

Query Match 30.2%; Score 118; DB 5; Length 561;
 Best Local Similarity 36.8%; Pred. No. 2e-07;
 Matches 21; Conservative 11; Mismatches 23; Indels 2; Gaps 2;

QY 12 CPSENEISRCDCRC-QRFPCNV-VKPLCIKICAPGCVCLRLGLRNKKKVCVPRSKC 66
 DB 123 CPVNEVSNECHNPTCKKCPQKNAPQVNCIMACQVCSMDGFEVRNNGVCVKEAEC 179

RESULT 3

Q9U1T5 PRELIMINARY; PRT; 626 AA.
 AC Q9U1T5;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Y69H2.3b protein.
 DE Y69H2.3B.
 GN Caenorhabditis elegans.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.A.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RT Science 282:2012-2018(1998).

DR EMBL; 298877; CAB54473.1; -.

DR HSSP; P56682; ICCV.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003645; FOLN.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR Pfam; PF01826; TIL; 4.
 DR SMART; SM00274; FOLN; 2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
 SQ SEQUENCE 626 AA; 67881 MW; AL7F7B445800E118 CRC64;

Query Match 30.2%; Score 118; DB 5; Length 626;
 Best Local Similarity 36.8%; Pred. No. 2.3e-07;
 Matches 21; Conservative 11; Mismatches 23; Indels 2; Gaps 2;

QY 12 CPSENEISRCDCRC-QRFPCNV-VKPLCIKICAPGCVCLRLGLRNKKKVCVPRSKC 66
 DB 123 CPVNEVSNECHNPTCKKCPQKNAPQVNCIMACQVCSMDGFEVRNNGVCVKEAEC 179

RESULT 4

O08523 PRELIMINARY; PRT; 2155 AA.
 ID O08523
 AC O08523;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Alpha tectorin.
 DE TECTA.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD1; TISSUE=COCHLEA;
 RX MEDLINE=97236843; PubMed=9079715;
 RA Legan P.K., Rau A., Keene J.N., Richardson G.P.;
 RT "Tectorins: modular matrix proteins of the inner ear homologous to
 RT components of the sperm-egg adhesion system.";
 RL J. Biol. Chem. 272:8791-8801(1997).
 DR EMBL; X99805; CA668138.1; -.
 DR HSSP; P00736; IAPQ.
 DR MGD; MGI:109575; Tecta.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001507; Endoglin/CD105.
 DR InterPro; IPR000421; FA58_C.
 DR InterPro; IPR003886; Nidogen_ext.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR InterPro; IPR001007; VWF_C.
 DR InterPro; IPR001846; VWF_D.
 DR Pfam; PF01826; TIL; 3.
 DR Pfam; PF00094; wvd; 4.
 DR Pfam; PF00100; zona_pellucida; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00539; NIDO; 1.
 DR SMART; SM00214; VWC; 1.
 DR SMART; SM00216; VWD; 4.
 DR SMART; SM00241; ZP; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS01286; FA58C_2; UNKNOWN_1.
 DR PROSITE; PS00682; ZP_DOMAIN; 1.
 KW EGF-like domain; Glycoprotein.
 SQ SEQUENCE 2155 AA; 239535 MW; 86C5C09AA392B1C5 CRC64;

Query Match 28.9%; Score 113; DB 11; Length 2155;
 Best Local Similarity 38.6%; Pred. No. 3.4e-06;
 Matches 22; Conservative 7; Mismatches 26; Indels 2; Gaps 2;

QY 11 KPSENEISRCDCRC-QRFPCNV-VKPLCIKICAPGCVCLRLGLRNKKKVCVPRSKC 67
 DB 983 ECPENSHFEEC-MTCTETCTETALGPICVDSCECCQDEGY-RLQSQCVTRSECG 1037

```

RESULT 5
Q99ND0 PRELIMINARY; PRT; 5374 AA.
ID AC Q99ND0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE ZAN (Zonadhesin).
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=21138439; PubMed=11239002;
RA Wilson M.D., Riener C., Martindale D.W., Schnupf P., Boright A.P.,
RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA Miller W., Koop B.F.;
RT "Comparative analysis of the gene-dense ACHE/TPR2 region on human
RT chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL Nucleic Acids Res. 29:1352-1365(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Cheung T.L., Wilson M.D., Koop B.F., Hardy D.M.;
RT "Genomic Basis of Inter- and Intra-species Variation in Zonadhesin
RT Domain Structure.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
DR EMBL; AF312033; AAK28824.1; -.
DR EMBL; AY046036; AAL04416.1; -.
DR MGD; MGI:106656; Zan.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR003645; FOLN.
DR InterPro; IPR000998; MAM domain.
DR InterPro; IPR003328; Tila_Cysrich.
DR InterPro; IPR002919; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00629; MAM; 3.
DR Pfam; PF01826; TIL; 25.
DR Pfam; PF02345; TILA; 25.
DR SMART; SM00094; vwd; 4.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00274; FOLN; 21.
DR SMART; SM00137; MAM; 3.
DR SMART; SM00214; VWC; 25.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 18.
DR PROSITE; PS50060; MAM_2; 3.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 5374 AA; 579536 MW; 90D2D8CFE5DE24EB CRC64;

Query Match 28.9%; Score 113; DB 11; Length 5374;
Best Local Similarity 37.5%; Pred. No. 8e-06;
Matches 24; Conservative 8; Mismatches 22; Indels 10; Gaps 3;

QY 11 KCPSNEIFSRCDGRCQPCNVVVKPLC-----IKICAPGCVCLGLYLRNKKKVCVPR 63
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 4741 KCPANSLYTHCLPTCLPSCSN--PDGRCGTSKAPSTCRGGCVCPGYYLLN-KDTCVHK 4797
::||
QY 64 SKCG 67
Db 4798 NQCG 4801

```

```

RESULT 6
O16488 PRELIMINARY; PRT; 166 AA.
ID AC O16488;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE Hypothetical 18.0 kDa protein.
GN B0238.12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Blanchard M., Bradshaw H.;
RT "The sequence of C. elegans cosmid B0238.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL; AF016450; AAB65990.1; -.
DR HSSP; P56682; ICCV.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
KW Hypothetical protein_2.
SQ SEQUENCE 166 AA; 17971 MW; DA92F2473442ADB CRC64;

Query Match 28.5%; Score 111.5; DB 5; Length 166;
Best Local Similarity 36.8%; Pred. No. 4.8e-07;
Matches 25; Conservative 11; Mismatches 25; Indels 7; Gaps 4;

QY 4 GGLGGRCCKCP---SNEIFSRCDGRCQPCNVVVKPLCIIKICAPG-CVCLGLYLRNKK-KK 58
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 27 GQVGGGQRLPCRGGRNEEYKTCGTACEPCTN--PNPMTKQCINNVCQCRSGYVRNEITR 84
||::|
QY 59 VCVPRSKC 66
Db 85 QCVPRQAC 92

```

```

RESULT 7
P90956 PRELIMINARY; PRT; 949 AA.
ID AC P90956;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE T01D3.3 protein.
GN T01D3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC Steward C.;
RT Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

```



```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Cheung T.L., Wassler M.J., Cornwall G.A., Hardy D.M.;
RT "Multiple Intra-species Variants of Human Zonadhesin.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
DR EMBL; AF332976; AAK01432.1; -.
DR HSSP; P56682; 1CCV.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000998; MAM_domain.
DR InterPro; IPR002965; P_rich_extensn.
DR InterPro; IPR003328; TILA_Cysrich.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF01826; TIL; 5.
DR Pfam; PF02345; TILA; 4.
DR Pfam; PF00094; vwd; 4.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00137; MAM; 3.
DR SMART; SM00216; VMD; 4.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00060; MAM_2; 3.
KW Glycoprotein.
SQ SEQUENCE 2689 AA; 7E12C3343BFD408C CRC64;

Query Match 27.9%; Score 109; DB 4; Length 2689;
Best Local Similarity 35.3%; Pred. No. 1.4e-05;
Matches 24; Conservative 11; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPSNEIFSRC-----DGRQRCPCPNVVPKPLCIKICAPGCVCRGLGYLRNKKV 59
Db 2210 ECPAYSSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICQPGYVLSEDK- 2260

QY 60 CVPRSKCG 67
Db 2261 CVPRSCG 2268

RESULT 14
Q96L85 PRELIMINARY; PRT; 2689 AA.
ID Q96L85
AC Q96L85;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Zonadhesin splice variant 2.
GN ZAN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheung T.L., Wilson M.D., Koop B.F., Hardy D.M.;
RT "Genomic Basis of Inter- and Intra-species Variation in Zonadhesin Domain Structure.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
DR EMBL; AY046055; AAL04415.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000998; MAM_domain.
DR InterPro; IPR003328; TILA_Cysrich.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00629; MAM; 3.
DR Pfam; PF01826; TIL; 5.
DR Pfam; PF02345; TILA; 4.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.

```

```

DR PROSITE; PS00740; MAM_1; UNKNOWN_1.
DR PROSITE; PS00060; MAM_2; 3.
KW Glycoprotein.
SQ SEQUENCE 2689 AA; 291277 MW; 2253CC9E22D7C45F CRC64;

Query Match 27.9%; Score 109; DB 4; Length 2689;
Best Local Similarity 35.3%; Pred. No. 1.4e-05;
Matches 24; Conservative 11; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPSNEIFSRC-----DGRQRCPCPNVVPKPLCIKICAPGCVCRGLGYLRNKKV 59
Db 2210 ECPAYSSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICQPGYVLSEDK- 2260

QY 60 CVPRSKCG 67
Db 2261 CVPRSCG 2268

RESULT 15
Q9BZ83 PRELIMINARY; PRT; 2721 AA.
ID Q9BZ83
AC Q9BZ83;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Zonadhesin variant 6.
GN ZAN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheung T.L., Wassler M.J., Cornwall G.A., Hardy D.M.;
RT "Multiple Intra-species Variants of Human Zonadhesin.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
DR EMBL; AF332980; AAK01436.1; -.
DR HSSP; P56682; 1CCV.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000998; MAM_domain.
DR InterPro; IPR002965; P_rich_extensn.
DR InterPro; IPR003328; TILA_Cysrich.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00629; MAM; 3.
DR Pfam; PF01826; TIL; 4.
DR Pfam; PF02345; TILA; 4.
DR Pfam; PF00094; vwd; 4.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00137; MAM; 3.
DR SMART; SM00216; VMD; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_4.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00060; MAM_2; 3.
KW Glycoprotein.
SQ SEQUENCE 2721 AA; 295506 MW; 142E9CBF3D404EC3 CRC64;

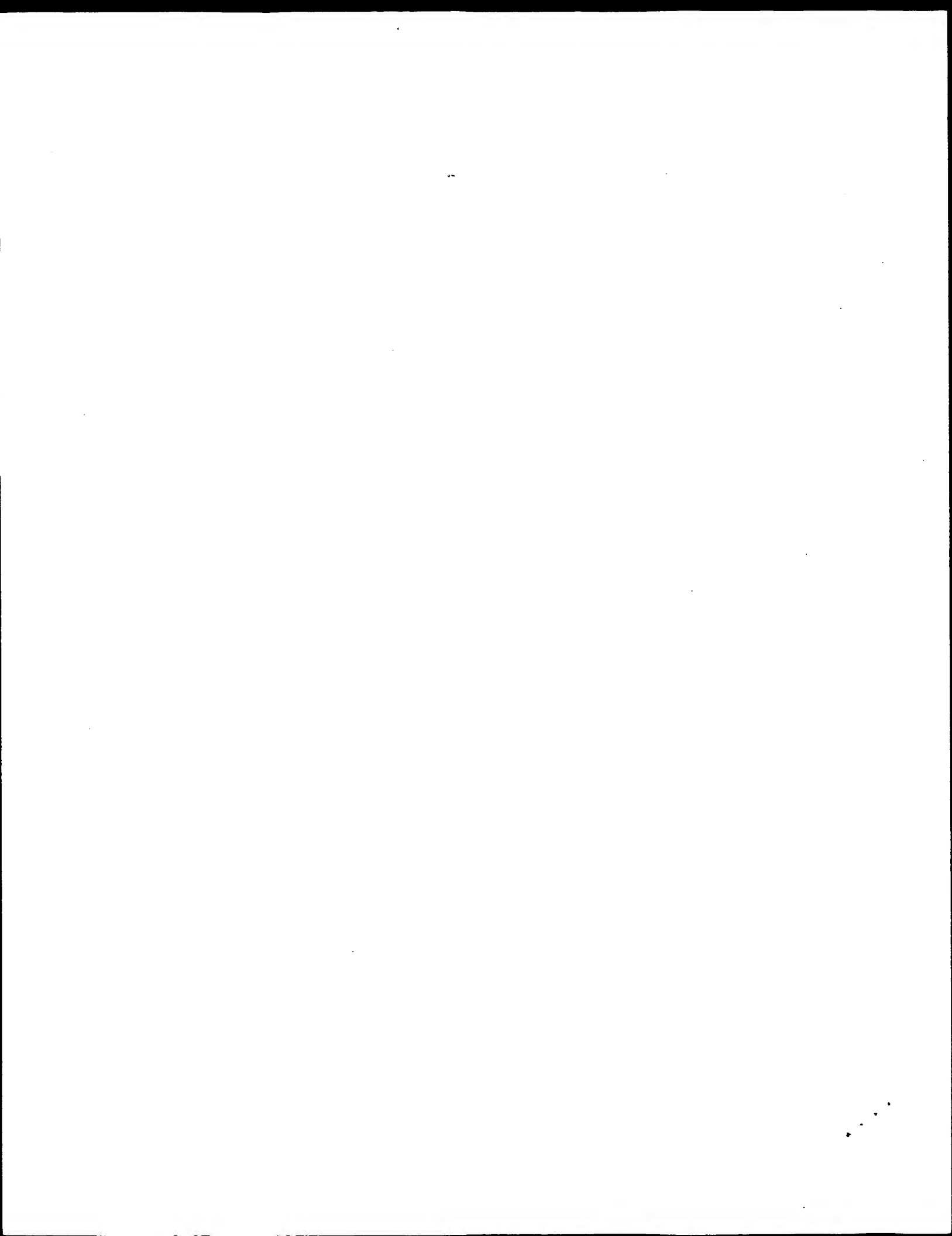
Query Match 27.9%; Score 109; DB 4; Length 2721;
Best Local Similarity 35.3%; Pred. No. 1.4e-05;
Matches 24; Conservative 11; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPSNEIFSRC-----DGRQRCPCPNVVPKPLCIKICAPGCVCRGLGYLRNKKV 59
Db 2210 ECPAYSSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICQPGYVLSEDK- 2260

QY 60 CVPRSKCG 67
Db 2261 CVPRSCG 2268

```

Search completed: February 26, 2003, 15:05:43
Job time : 31 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:03:23 ; Search time 36 seconds
(without alignments)
247.994 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGGGLGGGKCPSEIFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	391	100.0	67	ABB08330	Bee venom protein
2	391	100.0	69	ABB08331	Bee venom protein
3	391	100.0	71	ABB08332	Bee venom protein
4	391	100.0	73	ABB08333	Bee venom protein
5	391	100.0	92	AA199209	Amino acid sequenc
6	108	27.6	3843	AB071529	Drosophila melanog
7	107.5	27.5	2476	AA067738	Pig p105 zona pell
8	106.5	27.2	2594	AA014748	IgG-Fc binding pro
9	106.5	27.2	2957	ABG22214	Novel human diagno
10	106.5	27.2	5405	AA014749	IgG-Fc binding pro

11	101.5	26.0	7337	22	ABG22216	Novel human diagno
12	100.5	25.7	735	22	ABG22215	Novel human diagno
13	100	25.6	84	20	AA30432	Mature nematode ex
14	100	25.6	84	21	AA015317	A. caninum nematod
15	100	25.6	91	17	AA091701	AcenAPc2. Ancylos
16	100	25.6	91	20	AA30393	Nematode extracted
17	100	25.6	91	20	AA30454	Nematode extracted
18	100	25.6	91	21	AA015346	A. caninum nematod
19	98.5	25.2	508	22	ABG22213	Novel human diagno
20	98	25.1	115	20	AA222170	Drosophila Acp62F
21	98	25.1	115	22	AB057966	Drosophila melanog
22	98	25.1	741	17	AA016462	Human von Willebra
23	98	25.1	2813	7	AA060053	Sequence of von Wi
24	98	25.1	2813	7	AA060462	Sequence of human
25	98	25.1	2813	23	AA075317	Human von Willebra
26	98	25.1	2814	22	AA029530	Novel human secret
27	96	24.6	164	22	AA04265	Human gene 8 encod
28	96	24.6	164	22	AA001650	Human gene 19 enco
29	96	24.6	189	22	AA04268	Human gene 8 encod
30	96	24.6	191	22	AA04266	Human gene 8 encod
31	95.5	24.4	149	22	AA04263	Human gene 8 encod
32	92.5	23.7	108	22	AB063313	Drosophila melanog
33	92	23.5	701	22	AB06725	Drosophila melanog
34	91	23.3	82	20	AA30399	Nematode extracted
35	91	23.3	82	20	AA30422	Mature nematode ex
36	91	23.3	82	21	AA015293	A. ceylanicum nema
37	91	23.3	82	21	AA015307	A. ceylanicum nema
38	91	23.3	171	20	AA30435	Mature nematode ex
39	91	23.3	190	17	AA091710	AcenAP4. Ancylost
40	91	23.3	190	20	AA30384	Nematode extracted
41	88	22.5	2009	22	AB064069	Drosophila melanog
42	87.5	22.4	178	22	AB061389	Drosophila melanog
43	87.5	22.4	1438	22	ABG03631	Novel human diagno
44	87.5	22.4	1438	22	ABG28308	Novel human diagno
45	87.5	22.4	1439	22	ABG14849	Novel human diagno

ALIGNMENTS

RESULT 1
ABB08330
ID ABB08330 standard; protein: 67 AA.
XX
AC ABB08330;
XX
DT 18-JUN-2002 (first entry)
XX
DE Bee venom protein Api m 6.01.
XX
KW Bee venom; isoform: immunosuppressant; vaccine; Api m 6; immune response;
KW bee venom hypersensitivity; antibody; protein purification; Api m 6.01;
KW immunotherapy; allergen.
XX
OS Apis sp.
XX
PN WO2001188085-A2.
XX
PD 22-NOV-2001.
XX
PF 16-FEB-2001; 2001WO-1801736.
XX
PR 18-FEB-2000; 2000US-0506978.
XX
PA (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.
XX
PI Spertini F;
XX
DR WPI; 2002-082988/11.
XX
PT New bee venom polypeptides, useful for modulating immune responses e.g.
PT in individual hypersensitive to the venom and for identifying
PT individual at risk for bee venom hypersensitivity

XX PS Example 2; Page 26; 32pp; English.

XX CC The present sequence is that of one the four isoforms of Api m 6,

CC CC designated Api m 6.01. The sequence represents the central amino acid

CC CC sequence shared by all four isoforms (see ABB08331, ABB08332 and

CC CC ABB08333). The specification describes a substantially pure polypeptide,

CC CC Api m 6, derived from bee venom and found in four isoforms. The proteins

CC CC of the invention have immunosuppressant activity and may form the basis

CC CC as an allergen for immunotherapy. The protein is useful for identifying

CC CC an individual at risk for bee venom hypersensitivity. The method

CC CC comprises administering Api m 6 to the individual and measuring an immune

CC CC response raised, where a detectable immune response indicates that the

CC CC individual is at risk for bee venom hypersensitivity. Antibodies specific

CC CC for Api m 6 are useful for purifying the protein.

XX SQ Sequence 67 AA;

Query Match 100.0%; Score 391; DB 23; Length 67;

Best Local Similarity 100.0%; Pred. No. 8e-32;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Db 1 GFGGLGGRGKCPSEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

QY 61 VPRSKCG 67

Db 61 VPRSKCG 67

RESULT 2

ABB08331

ID ABB08331 standard; protein; 69 AA.

XX AC ABB08331;

XX DT 18-JUN-2002 (first entry)

XX DE Bee venom protein Api m 6.02.

XX KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;

XX KW bee venom hypersensitivity; antibody; protein purification; Api m 6.02;

XX KW immunotherapy; allergen.

XX OS Apis sp.

XX PN W0200188085-A2.

XX PD 22-NOV-2001.

XX PF 16-FEB-2001; 2001WO-IB01736.

XX PR 18-FEB-2000; 2000US-0506978.

XX PA (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

XX PI Spertini F;

XX DR WPI; 2002-082988/11.

XX CC New bee venom polypeptides, useful for modulating immune responses e.g.

XX CC in individual hypersensitive to the venom and for identifying

XX CC individual at risk for bee venom hypersensitivity

XX CC Example 2; Page 26; 32pp; English.

XX CC The present sequence is that of one the four isoforms of Api m 6,

CC CC designated Api m 6.02. All four isoforms share a common central amino

CC CC acid sequence shared by all four isoforms (see ABB08330, ABB08332 and

CC CC ABB08333). The specification describes a substantially pure polypeptide,

CC CC Api m 6, derived from bee venom and found in four isoforms. The proteins

CC CC of the invention have immunosuppressant activity and may form the basis

CC CC as a vaccine. Api m 6 is useful for modulating an immune response, i.e.

CC CC as an allergen for immunotherapy. The protein is useful for identifying

CC CC an individual at risk for bee venom hypersensitivity. The method

CC CC comprises administering Api m 6 to the individual and measuring an immune

CC CC response raised, where a detectable immune response indicates that the

CC CC individual is at risk for bee venom hypersensitivity. Antibodies specific

CC CC for Api m 6 are useful for purifying the protein.

XX SQ Sequence 69 AA;

Query Match 100.0%; Score 391; DB 23; Length 69;

Best Local Similarity 100.0%; Pred. No. 8.2e-32;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Db 1 GFGGLGGRGKCPSEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

QY 61 VPRSKCG 67

Db 61 VPRSKCG 67

RESULT 3

ABB08332

ID ABB08332 standard; protein; 71 AA.

XX AC ABB08332;

XX DT 18-JUN-2002 (first entry)

XX DE Bee venom protein Api m 6.03.

XX KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;

XX KW bee venom hypersensitivity; antibody; protein purification; Api m 6.03;

XX KW immunotherapy; allergen.

XX OS Apis sp.

XX PN W0200188085-A2.

XX PD 22-NOV-2001.

XX PF 16-FEB-2001; 2001WO-IB01736.

XX PR 18-FEB-2000; 2000US-0506978.

XX PA (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

XX PI Spertini F;

XX DR WPI; 2002-082988/11.

XX CC New bee venom polypeptides, useful for modulating immune responses e.g.

XX CC in individual hypersensitive to the venom and for identifying

XX CC individual at risk for bee venom hypersensitivity

XX CC Example 2; Page 26; 32pp; English.

XX CC The present sequence is that of one the four isoforms of Api m 6,

CC CC designated Api m 6.03. All four isoforms share a common central amino

CC CC acid sequence shared by all four isoforms (see ABB08330, ABB08331 and

CC CC ABB08333). The specification describes a substantially pure polypeptide,

CC CC Api m 6, derived from bee venom and found in four isoforms. The proteins

CC CC of the invention have immunosuppressant activity and may form the basis

CC CC as a vaccine. Api m 6 is useful for modulating an immune response, i.e.

CC CC as an allergen for immunotherapy. The protein is useful for identifying

CC CC an individual at risk for bee venom hypersensitivity. The method

CC CC comprises administering Api m 6 to the individual and measuring an immune

CC CC response raised, where a detectable immune response indicates that the

CC CC individual is at risk for bee venom hypersensitivity. Antibodies specific

CC CC for Api m 6 are useful for purifying the protein.

CC as an allergen for immunotherapy. The protein is useful for identifying
 CC an individual at risk for bee venom hypersensitivity. The method
 CC comprises administering Api m 6 to the individual and measuring an immune
 CC response raised, where a detectable immune response indicates that the
 CC individual is at risk for bee venom hypersensitivity. Antibodies specific
 CC for Abi m 6 are useful for purifying the protein.

XX SQ Sequence 71 AA;

Query Match 100.0%; Score 391; DB 23; Length 71;
 Best Local Similarity 100.0%; Pred. No. 8.4e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGGLGGKCPSPNEIFSRDGRCPNVVVKPLCIKICAPGCVCLGLYLRNKKKVC 60
 |||||

Db 5 GGFGGLGGKCPSPNEIFSRDGRCPNVVVKPLCIKICAPGCVCLGLYLRNKKKVC 64

QY 61 VPRSKCG 67
 |||||

Db 65 VPRSKCG 71

RESULT 4
 ABB08333
 ID ABB08333 standard; protein; 73 AA.

XX AC ABB08333;

XX DT 18-JUN-2002 (first entry)

XX DE Bee venom protein Api m 6.04.

XX KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;
 KW bee venom hypersensitivity; antibody; protein purification; Api m 6.04;
 KW immunotherapy; allergen.

XX OS Apis sp.

XX FH Key Location/Qualifiers

XX FT Misc-difference 72 /label= Pro, Leu

XX FT Misc-difference 73 /label= Pro, Leu

XX PN WO2001188085-A2.

XX PD 22-NOV-2001.

XX PF 16-FEB-2001; 2001WO-IB01736.

XX PR 18-FEB-2000; 2000US-0506978.

XX PA (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

XX PI Spertini F;

XX DR WPI; 2002-082988/11.

XX PT New bee venom polypeptides, useful for modulating immune responses e.g.
 PT in individual hypersensitive to the venom and for identifying
 PT individual at risk for bee venom hypersensitivity

XX PS Example 2; Page 26; 32pp; English.

XX CC The present sequence is that of one the four isoforms of Api m 6,
 CC designated Api m 6.04. All four isoforms share a common central amino
 CC acid sequence shared by all four isoforms (see ABB08330, ABB08331 and
 CC ABB08332). The specification describes a substantially pure polypeptide,
 CC Api m 6, derived from bee venom and found in four isoforms. The proteins
 CC of the invention have immunosuppressant activity and may form the basis
 CC of a vaccine. Api m 6 is useful for modulating an immune response, i.e.
 CC as an allergen for immunotherapy. The protein is useful for identifying
 CC an individual at risk for bee venom hypersensitivity. The method

CC comprises administering Api m 6 to the individual and measuring an immune
 CC response raised, where a detectable immune response indicates that the
 CC individual is at risk for bee venom hypersensitivity. Antibodies specific
 CC for Abi m 6 are useful for purifying the protein.

XX SQ Sequence 73 AA;

Query Match 100.0%; Score 391; DB 23; Length 73;
 Best Local Similarity 100.0%; Pred. No. 8.6e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGGLGGKCPSPNEIFSRDGRCPNVVVKPLCIKICAPGCVCLGLYLRNKKKVC 60
 |||||

Db 5 GGFGGLGGKCPSPNEIFSRDGRCPNVVVKPLCIKICAPGCVCLGLYLRNKKKVC 64

QY 61 VPRSKCG 67
 |||||

Db 65 VPRSKCG 71

RESULT 5
 AAY69209
 ID AAY69209 standard; Protein; 92 AA.

XX AC AAY69209;

XX DT 30-MAY-2000 (first entry)

XX DE Amino acid sequence of honey bee venom PX3.101 protein.

XX KW Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor;
 KW CXCR1; CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
 KW inflammatory disease; gene therapy; cancer; autoimmune disease; pain;
 KW chemokine imbalance; rheumatoid arthritis; multiple sclerosis;
 KW psoriasis; systemic lupus erythematosus; Crohn's disease; vasculitis;
 KW scleroderma; metastatic cancer; Alzheimer's disease; wound healing;
 KW aging process; antigen.

XX OS Apis mellifera.

XX FH Key Location/Qualifiers

XX FT Peptide 1..19 /note= "signal peptide"

XX FT Region 20..34 /note= "this region contains 5 GGX repeats"

XX PN GB2341389-A.

XX PD 15-MAR-2000.

XX PF 13-SEP-1999; 99GB-0021605.

XX PR 14-SEP-1998; 98US-0100172.

XX PA (PANP-) PAN PACIFIC PHARM INC.

XX PI Chi X, Lu Y;

XX DR WPI; 2000-185368/17.

XX DR N-PSDB; AAZ61247.

XX PT Isolated nucleic acids encoding the bee venom protein PX3.101 useful
 PT for treating autoimmune and inflammatory disorders such as rheumatoid
 PT arthritis

XX PS Claim 2; Fig 3A-B; 83pp; English.

XX CC The present sequence represents the protein PX3.101, which is a honey
 CC bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
 CC interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and
 CC inhibits a variety of enzymes (e.g. cyclooxygenases, lipoxigenases,
 CC phospholipases and proteases) associated with inflammatory diseases.
 CC The nucleic acids may be used for the recombinant production of


```

RESULT 8
AAW14748
ID AAW14748 standard; Protein; 2594 AA.
XX
AC AAW14748;
XX
DT 13-MAY-1997 (first entry)
XX
DE IgG-Fc binding protein encoded by 7.8 kb fragment of pNV11-ST.
XX
KW Fragment 13; pNV11-ST; IgG-Fc binding protein; immunoglobulin; K17;
KW human; colonic epithelium; monoclonal antibody; K9; probe.
XX
OS Homo sapiens.
XX
PN WO9527057-A1.
XX
PD 12-OCT-1995.
XX
PF 03-APR-1995; 95WO-JP00638.
XX
PR 30-MAR-1995; 95JP-0109927.
PR 01-APR-1994; 94JP-0129487.
PR 24-AUG-1994; 94JP-0222547.
XX
PA (CHUS ) CHUGAI SEIYAKU KK.
XX
PI Harada N, Morikawa M;
XX
WPI; 1995-358632/46.
DR N-PSDB; AAT63073.
XX
DNA derived from colonic epithelium encoding IgG-Fc binding protein
PT - used in the mapping and analysis of IgG-Fc binding protein mRNA
XX
PS Claim 1; Page 71-84; 132pp; Japanese.
XX
CC This sequence is encoded by fragment 13 which is a NotI/KpnI fragment
CC from pNV11-ST. This sequence represents a portion of the IgG-Fc binding
CC protein of human colonic epithelium. This sequence was used in the
CC isolation of the full length sequence given in AAW14749. mRNA isolated
CC from human colonic epithelial tissue was used to prepare a cDNA library.
CC This was screened using monoclonal antibodies K9 and K17 which bind to
CC the large and small components of the binding protein. Active clones,
CC see also AAT63077-81, were used to derive probes for screening a second
CC DNA library from human colonic epithelial tissue.
XX
SQ Sequence 2594 AA;
XX
Query Match 27.2%; Score 106.5; DB 16; Length 2594;
Best Local Similarity 36.8%; Pred. No. 0.035;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
XX
QY 11 KCPSEIFSRCDGRCQFCFNNVVPKPLCIKICAPGCVCRGLYLRNKKVCPVRSKCG 67
:| | : | | : | | : | | : | | : | | : | | : | | : | |
Db 1531 ECPNPNHYELCADTCSLGCSALSAPPQCDGCAEGCQCDGFLYN-GQACVPIQOCG 1586

RESULT 9
ABG22214
ID ABG22214 standard; Protein; 2957 AA.
XX
AC ABG22214;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #22205.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX

```

```

OS Homo sapiens.
XX
PN WO200175067-A2;
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
DR N-PSDB; AAS86401.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
Claim 20; SEQ ID No 52573; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(CC (II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
imaging of sites expressing (II). (I) and (II) are useful in medical
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG00010-ABG0377 represent novel human
diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2957 AA;
XX
Query Match 27.2%; Score 106.5; DB 22; Length 2957;
Best Local Similarity 36.8%; Pred. No. 0.039;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
XX
QY 11 KCPSEIFSRCDGRCQFCFNNVVPKPLCIKICAPGCVCRGLYLRNKKVCPVRSKCG 67
:| | : | | : | | : | | : | | : | | : | | : | | : | |
Db 1623 ECPNPNHYELCADTCSLGCSALSAPPQCDGCAEGCQCDGFLYN-GQACVPIQOCG 1678

RESULT 10
AAW14749
ID AAW14749 standard; Protein; 5405 AA.
XX
AC AAW14749;
XX
DT 13-MAY-1997 (first entry)
XX
DE IgG-Fc binding protein.
XX
KW Fragment 13; pNV11-ST; IgG-Fc binding protein; immunoglobulin; K17;
KW human; colonic epithelium; monoclonal antibody; K9; probe.
XX
OS Homo sapiens.
XX

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:05:12 : Search time 33 seconds
(without alignments)
76.587 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGGLGGRCKPSENEFSR.....CRLGYLRNKKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	106.5	27.2	5405	9	US-10-025-380-1116
2	106.5	27.2	5405	10	US-09-922-217-1116
3	98	25.1	2813	10	US-09-381-261A-1
4	87.5	22.4	2813	10	US-09-886-900-2
5	85	21.7	759	9	US-10-189-971-22
6	85	21.7	1057	9	US-10-189-971-6
7	85	21.7	1192	9	US-10-189-971-18
8	85	21.7	1207	9	US-10-189-971-20
9	85	21.7	1251	9	US-10-189-971-16
10	85	21.7	1342	9	US-10-189-971-24
11	85	21.7	1477	9	US-10-189-971-8
12	85	21.7	1512	9	US-10-189-971-10
13	85	21.7	1535	9	US-10-189-971-14
14	85	21.7	1570	9	US-10-189-971-12
15	85	21.7	1593	9	US-10-189-971-4
16	85	21.7	1628	9	US-10-189-971-2
17	83.5	21.4	355	10	US-09-826-212-14
18	83.5	21.4	355	10	US-09-935-727-16
19	82.5	21.1	469	10	US-09-925-301-1279

20	78.5	20.1	77	9	US-09-950-933A-46	Sequence 46, Appl
21	77.5	19.8	84	9	US-09-950-933A-54	Sequence 54, Appl
22	76.5	19.6	855	10	US-09-900-751-2	Sequence 2, Appl
23	75.5	19.3	349	10	US-09-826-212-13	Sequence 13, Appl
24	75.5	19.3	349	10	US-09-935-727-15	Sequence 15, Appl
25	75.5	19.3	5179	9	US-10-025-380-1068	Sequence 1068, Ap
26	75.5	19.3	5179	10	US-09-922-217-1068	Sequence 1068, Ap
27	75.5	19.3	5179	10	US-09-833-263-1068	Sequence 1068, Ap
28	72	18.4	366	10	US-09-205-658-103	Sequence 103, App
29	72	18.4	366	10	US-09-844-353A-103	Sequence 103, App
30	72	18.4	1367	9	US-09-870-759-120	Sequence 120, App
31	71	18.2	129	9	US-09-950-933A-43	Sequence 43, Appl
32	70	17.9	474	9	US-10-084-994-11	Sequence 11, Appl
33	68.5	17.5	3046	9	US-09-759-130B-441	Sequence 441, App
34	67	17.1	578	10	US-09-908-322-13	Sequence 13, App
35	66	16.9	1050	9	US-09-796-753-114	Sequence 114, App
36	66	16.9	1509	10	US-09-901-940-2	Sequence 2, Appli
37	65.5	16.8	96	9	US-09-950-933A-60	Sequence 60, Appli
38	65.5	16.8	689	10	US-09-071-838-2	Sequence 2, Appli
39	65	16.6	636	9	US-09-796-753-100	Sequence 100, App
40	65	16.6	636	9	US-09-796-753-124	Sequence 124, App
41	64	16.4	276	9	US-10-086-176A-5	Sequence 5, Appli
42	64	16.4	276	9	US-09-741-106-9	Sequence 9, Appli
43	64	16.4	276	10	US-09-766-778-1	Sequence 1, Appli
44	64	16.4	291	9	US-09-924-340-48	Sequence 48, Appl
45	64	16.4	291	9	US-09-924-340-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-10-025-380-1116
: Sequence 1116, Application US/10025380
: Publication No. US20020182191A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Lodes, Michael J.
: APPLICANT: Secrist, Heather
: APPLICANT: Benson, Darin R.
: APPLICANT: Meagher, Madeleine Joy
: APPLICANT: Stolk, John A.
: APPLICANT: Wang, Tongtong
: APPLICANT: Jiang, Yugu
: APPLICANT: Smith, Carole L.
: APPLICANT: King, Gordon E.
: APPLICANT: Wang, Aijun
: APPLICANT: Clapper, Jonathan D.
: APPLICANT: Skeiky, Yasir A. W.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedvick Thomas S.
: APPLICANT: Carter, Darrick
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
: FILE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
: FILE REFERENCE: 210121.471C14
: CURRENT APPLICATION NUMBER: US/10/025,380
: CURRENT FILING DATE: 2001-12-19
: NUMBER OF SEQ ID NOS: 1129
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1116
: LENGTH: 5405
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-025-380-1116

Query Match 27.2%; Score 106.5; DB 9; Length 5405;
Best Local Similarity 36.8%; Pred. No. 0.027;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
QY 11 KCPSEIFSRCDGRCRCPCNNVVKPLKIKICAPCCVRLGRLNKKKVCVPRSKCG 67
DB 1531 ECPNHNVELCADTCLGCSALSAPPQCDGCAEGCQDSGLYN-GQACVPIQCCG 1586

Db 695 MDERGDCVPKAC 707

RESULT 4

US-09-886-900-2

Sequence 2, Application US/09886900

Patent No. US20020137051A1

GENERAL INFORMATION:

APPLICANT: Venta, Patrick J

Yuzbasiyan-Gurkan, Vilma

Schall, William D

Brewer, George J

TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND

FACTOR AND METHODS OF USE

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: 3445 Corporate Drive

CITY: Troy

STATE: Michigan

COUNTRY: USA

ZIP: 48098

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/886,900

FILING DATE: 21-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/896,449

FILING DATE: 18-JUL-1997

ATTORNEY/AGENT INFORMATION:

NAME: Smith, DeAnn F.

REFERENCE/DOCKET NUMBER: 2115-001226

TELECOMMUNICATION INFORMATION:

TELEPHONE: 248-641-1600

TELEFAX: 248-641-0270

TELEX: 287637

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2813 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-886-900-2

Query Match 22.4%; Score 87.5; DB 10; Length 2813;

Best Local Similarity 31.0%; Pred. No. 0.99;

Matches 18; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

QY 9 RCKPFSNIFSRCDGRCQFCPNVVPKPLCTKICAPGCVRLGVLNRKKKVCVPRSKC 66

Db 292 RPACFAGMEYKECVSPCTRTQSLHVKVQCQVDCGSCPEGLLDREGH-CVGSAC 348

RESULT 5

US-10-189-971-22

Sequence 22, Application US/10189971

Publication No. US20030028907A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Scoville, John

APPLICANT: Turner, C. Alexander Jr.

TITLE OF INVENTION: NO. US20030028907A1 Human Kiellin-like Proteins and Polynucle

FILE OF INVENTION: Same

FILE REFERENCE: LEX-0360-USA

CURRENT APPLICATION NUMBER: US/10/189,971

CURRENT FILING DATE: 2002-07-03

PRIOR APPLICATION NUMBER: US 60/302,949

Db 695 MDERGDCVPKAC 707

RESULT 2

US-09-922-217-1116

Sequence 1116, Application US/09922217

Patent No. US20020076414A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Lodes, Michael J.

APPLICANT: Secrist, Heather

APPLICANT: Benson, Darin R.

APPLICANT: Meagher, Madeleine Joy

APPLICANT: Stolk, John A.

APPLICANT: Wang, Tongtong

APPLICANT: Jiang, Yudi

APPLICANT: Smith, Carole Lynn

APPLICANT: King, Gordon E.

APPLICANT: Wang, Aijun

APPLICANT: Clapper, Jonathan D.

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.471C13

CURRENT APPLICATION NUMBER: US/09/922,217

CURRENT FILING DATE: 2001-08-03

NUMBER OF SEQ ID NOS: 1124

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1116

LENGTH: 5405

TYPE: PRT

ORGANISM: Homo sapiens

US-09-922-217-1116

Query Match 27.2%; Score 106.5; DB 10; Length 5405;

Best Local Similarity 36.8%; Pred. No. 0.027;

Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;

QY 11 KCPNFSRCDGRCQFCPNVVPKPLCTKICAPGCVRLGVLNRKKKVCVPRSKC 67

Db 1531 ECPNFSHYELCADTCSLGSALSAPPQCDGCAEGCQDSGFLYN-GQACVPIQQCG 1586

RESULT 3

US-09-381-261A-1

Sequence 1, Application US/09381261A

Patent No. US20020123457A1

GENERAL INFORMATION:

APPLICANT: Loscalzo, Joseph

APPLICANT: Inbal, Aida

TITLE OF INVENTION: NO. US20020123457A1 Anti-Platelet Agent

FILE REFERENCE: 102258.327

CURRENT APPLICATION NUMBER: US/09/381,261A

CURRENT FILING DATE: 1999-12-07

PRIOR APPLICATION NUMBER: PCT/US98/06092

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: US 60/046,981

PRIOR FILING DATE: 1997-03-27

NUMBER OF SEQ ID NOS: 3

SOFTWARE: Patentin version 3.1

SEQ ID NO 1

LENGTH: 2813

TYPE: PRT

ORGANISM: Homo sapiens

US-09-381-261A-1

Query Match 25.1%; Score 98; DB 10; Length 2813;

Best Local Similarity 26.0%; Pred. No. 0.098;

Matches 19; Conservative 13; Mismatches 27; Indels 14; Gaps 2;

QY 8 GRG-----KCPNFSRCDGRCQFCPNV-VPKPLCTKICAPGCVRLGVL 53

Db 635 GRGVRVAVRPFGRCELNCPGQVYLCGTPTCNLTCSLSYDDECNACLEGCFPPGLY 694

QY 54 RNKKKVCVPRSKC 66

; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: US 60/315,634
 ; PRIOR FILING DATE: 2001-08-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 22
 ; LENGTH: 759
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-10-189-971-22

Query Match 21.7%; Score 85; DB 9; Length 759;
 Best Local Similarity 30.6%; Pred. No. 0.55;
 Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCQRCFPCNV-VP-----KPLCIKICAPGCVCRGLGYLRNKKKVCVPRS 64
 Db 674 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCPAGLVEHEAH-CIPPE 732
 QY 65 KC 66
 Db 733 AC 734

RESULT 6

US-10-189-971-6
 ; Sequence 6, Application US/10189971
 ; Publication No. US20030028907A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Scoville, John
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. US20030028907A1el Human Kiellin-like Proteins and Polynucleotides
 ; FILE REFERENCE: LEX-0360-USA
 ; CURRENT APPLICATION NUMBER: US/10/189, 971
 ; CURRENT FILING DATE: 2002-07-03
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: US 60/302,949
 ; PRIOR FILING DATE: 2001-08-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1057
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-10-189-971-6

Query Match 21.7%; Score 85; DB 9; Length 1057;
 Best Local Similarity 30.6%; Pred. No. 0.73;
 Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCQRCFPCNV-VP-----KPLCIKICAPGCVCRGLGYLRNKKKVCVPRS 64
 Db 972 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCPAGLVEHEAH-CIPPE 1030
 QY 65 KC 66
 Db 1031 AC 1032

RESULT 7

US-10-189-971-18
 ; Sequence 18, Application US/10189971
 ; Publication No. US20030028907A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Scoville, John
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. US20030028907A1el Human Kiellin-like Proteins and Polynucleotides
 ; FILE REFERENCE: LEX-0360-USA

; CURRENT APPLICATION NUMBER: US/10/189,971
 ; CURRENT FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: US 60/302,949
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: US 60/315,634
 ; PRIOR FILING DATE: 2001-08-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 1192
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-10-189-971-18

Query Match 21.7%; Score 85; DB 9; Length 1192;
 Best Local Similarity 30.6%; Pred. No. 0.82;
 Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCQRCFPCNV-VP-----KPLCIKICAPGCVCRGLGYLRNKKKVCVPRS 64
 Db 1107 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCPAGLVEHEAH-CIPPE 1165
 QY 65 KC 66
 Db 1166 AC 1167

RESULT 8

US-10-189-971-20
 ; Sequence 20, Application US/10189971
 ; Publication No. US20030028907A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Scoville, John
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. US20030028907A1el Human Kiellin-like Proteins and Polynucleotides
 ; FILE REFERENCE: LEX-0360-USA
 ; CURRENT APPLICATION NUMBER: US/10/189, 971
 ; CURRENT FILING DATE: 2002-07-03
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: US 60/302,949
 ; PRIOR FILING DATE: 2001-08-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 1207
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-10-189-971-20

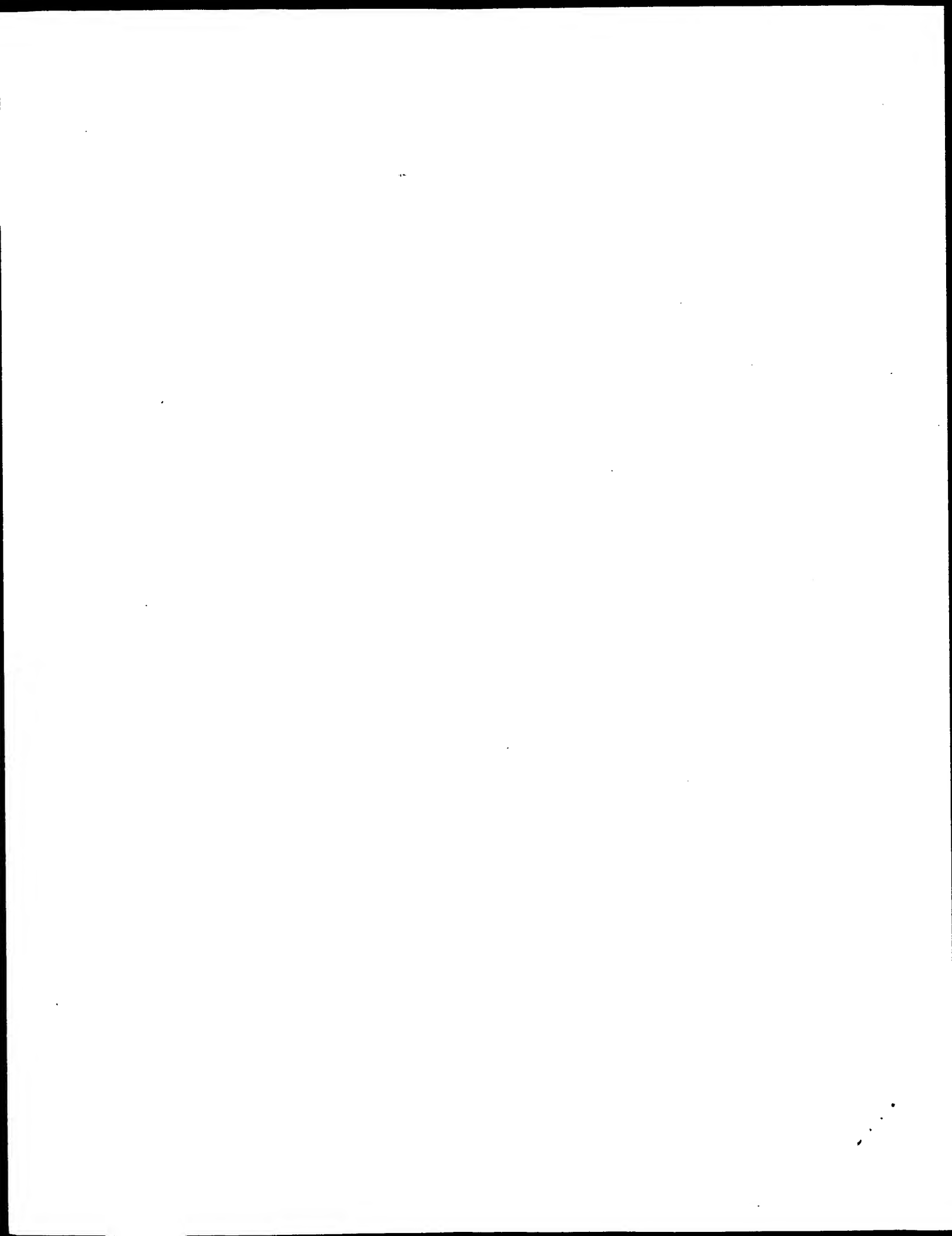
Query Match 21.7%; Score 85; DB 9; Length 1207;
 Best Local Similarity 30.6%; Pred. No. 0.82;
 Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCQRCFPCNV-VP-----KPLCIKICAPGCVCRGLGYLRNKKKVCVPRS 64
 Db 1122 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCPAGLVEHEAH-CIPPE 1180
 QY 65 KC 66
 Db 1181 AC 1182

RESULT 9

US-10-189-971-16
 ; Sequence 16, Application US/10189971
 ; Publication No. US20030028907A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Scoville, John
 ; APPLICANT: Turner, C. Alexander Jr.

[illegible]



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:04:07 : Search time 141 seconds
(without alignments)
306.363 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGRGKCPSEIFSR.....CRLGYLRNKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	391	100.0	67	19 US-09-506-978-1	Sequence 1, Appli
2	391	100.0	67	25 US-10-174-151-1	Sequence 1, Appli
3	391	100.0	67	26 US-10-204-145-1	Sequence 1, Appli
4	391	100.0	69	19 US-09-506-978-2	Sequence 2, Appli
5	391	100.0	69	25 US-10-174-151-2	Sequence 2, Appli
6	391	100.0	69	26 US-10-204-145-2	Sequence 2, Appli

7	391	100.0	71	19 US-09-506-978-3	Sequence 3, Appli
8	391	100.0	71	25 US-10-174-151-3	Sequence 3, Appli
9	391	100.0	71	26 US-10-204-145-3	Sequence 3, Appli
10	391	100.0	73	19 US-09-506-978-4	Sequence 4, Appli
11	391	100.0	73	25 US-10-174-151-4	Sequence 4, Appli
12	391	100.0	73	26 US-10-204-145-4	Sequence 4, Appli
13	391	100.0	92	1 PCT-US99-21077-2	Sequence 2, Appli
14	113.5	29.0	62	27 US-60-142-896-1491	Sequence 1491, Ap
15	113.5	29.0	62	27 US-60-145-138-701	Sequence 701, App
16	113.5	29.0	197	16 US-09-270-767-43953	Sequence 43953, A
17	111.5	28.5	166	27 US-60-360-039-6314	Sequence 6314, Ap
18	109	27.9	1537	27 US-60-230-445-1775	Sequence 1775, Ap
19	109	27.9	2554	27 US-60-230-445-1881	Sequence 1881, Ap
20	109	27.9	2593	27 US-60-207-583-584	Sequence 584, App
21	109	27.9	2601	24 US-10-016-248-59	Sequence 59, Appl
22	109	27.9	2659	27 US-60-207-583-632	Sequence 632, App
23	108	27.6	164	27 US-60-138-684-918	Sequence 918, App
24	108	27.6	1010	27 US-60-147-189-857	Sequence 857, App
25	108	27.6	3644	27 US-60-167-324-871	Sequence 871, App
26	108	27.6	3644	27 US-60-171-625-561	Sequence 561, App
27	108	27.6	3644	27 US-60-173-386-841	Sequence 841, App
28	108	27.6	3644	27 US-60-175-871-955	Sequence 955, App
29	108	27.6	3644	27 US-60-184-775-869	Sequence 869, App
30	108	27.6	3644	27 US-60-191-637-41010	Sequence 41010, A
31	108	27.6	3644	27 US-60-191-700-930	Sequence 930, App
32	108	27.6	3843	20 US-09-614-150-41379	Sequence 41379, A
33	106.5	27.2	76	27 US-60-192-739-2374	Sequence 2374, Ap
34	106.5	27.2	2957	1 PCT-US01-08631-52573	Sequence 52573, A
35	106.5	27.2	3014	27 US-60-258-273-160	Sequence 160, App
36	106.5	27.2	5405	23 US-09-922-217-1116	Sequence 1116, Ap
37	106.5	27.2	5405	24 US-10-025-380-1116	Sequence 1116, Ap
38	101.5	26.0	77	27 US-60-196-174-1252	Sequence 1252, Ap
39	101.5	26.0	91	27 US-60-162-243-571	Sequence 571, App
40	101.5	26.0	126	27 US-60-182-247-3514	Sequence 3514, Ap
41	101.5	26.0	126	27 US-60-177-571-4507	Sequence 4507, Ap
42	101.5	26.0	126	27 US-60-177-646-3718	Sequence 3718, Ap
43	101.5	26.0	7337	1 PCT-US01-08631-52575	Sequence 52575, A
44	100.5	25.7	135	27 US-60-196-174-1284	Sequence 1284, Ap
45	100.5	25.7	735	1 PCT-US01-08631-52574	Sequence 52574, A

ALIGNMENTS

RESULT 1

US-09-506-978-1

; Sequence 1, Application US/09506978

; GENERAL INFORMATION:

; APPLICANT: Spertini, Francois

; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF

; FILE REFERENCE: 18519-001

; CURRENT APPLICATION NUMBER: US/09/506,978

; CURRENT FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 67

; TYPE: PRT

; ORGANISM: Apis mellifera

US-09-506-978-1

Query Match 100.0%; Score 391; DB 19; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.le-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGRGKCPSEIFSRCDRCORFCPNVVPKLCIKICAPGCVCLGYLRNKKVC 60

Db 1 GGFGLGGRGKCPSEIFSRCDRCORFCPNVVPKLCIKICAPGCVCLGYLRNKKVC 60

QY 61 VPRSKCG 67

Db 61 VPRSKCG 67

```
RESULT 2
US-10-174-151-1
; Sequence 1, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-174-151-1
Query Match      100.0%; Score 391; DB 25; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 3
US-10-204-145-1
; Sequence 1, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-204-145-1
Query Match      100.0%; Score 391; DB 26; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 4
US-09-506-978-2
; Sequence 2, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
```

```
; CURRENT APPLICATION NUMBER: US/09/506,978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 68 and 69 is either PL or LP.
US-09-506-978-2
Query Match      100.0%; Score 391; DB 19; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 5
US-10-174-151-2
; Sequence 2, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 68 and 69 is either PL or LP.
US-10-174-151-2
Query Match      100.0%; Score 391; DB 25; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 6
US-10-204-145-2
; Sequence 2, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 2
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 68 and 69 is either PL or LP.
US-10-204-145-2

Query Match          100.0%; Score 391; DB 26; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 60
Db 1 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 7
US-09-506-978-3
; Sequence 3, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/09/506,978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Apis mellifera
US-09-506-978-3

Query Match          100.0%; Score 391; DB 19; Length 71;
Best Local Similarity 100.0%; Pred. No. 3.3e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 60
Db 5 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 64

Qy 61 VPRSKCG 67
Db 65 VPRSKCG 71

RESULT 8
US-10-174-151-3
; Sequence 3, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-174-151-3

Query Match          100.0%; Score 391; DB 25; Length 71;
Best Local Similarity 100.0%; Pred. No. 3.3e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 60
Db 5 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 64

Qy 61 VPRSKCG 67
Db 65 VPRSKCG 71

RESULT 9
US-10-204-145-3
; Sequence 3, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064, 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-204-145-3

Query Match          100.0%; Score 391; DB 26; Length 71;
Best Local Similarity 100.0%; Pred. No. 3.3e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 60
Db 5 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 64

Qy 61 VPRSKCG 67
Db 65 VPRSKCG 71

RESULT 10
US-09-506-978-4
; Sequence 4, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/09/506,978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
US-09-506-978-4

Query Match          100.0%; Score 391; DB 19; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 60
Db 5 GGFGLGGRGKCPSPNEIFSRCDGRQFCFNNVVPKPLCIKICAPGCVCLGLYLRNKKVC 64

Qy 61 VPRSKCG 67
Db 65 VPRSKCG 71

```

```

; Sequence 2, Application PC/TUS9921077
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yuefeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200PC
; CURRENT APPLICATION NUMBER: PCT/US99/21077
; CURRENT FILING DATE: 1999-09-13
; EARLIER APPLICATION NUMBER: US 60/100,172
; EARLIER FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Apis mellifera
; PCT-US99-21077-2

Query Match      100.0%; Score 391; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.1e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGLYLRNKKKVC 60
    |||||
DB 26 GFGGLGGRGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGLYLRNKKKVC 85

QY 61 VPRSKCG 67
    |||||
DB 86 VPRSKCG 92

RESULT 14
US-60-142-896-1491
; Sequence 1491, Application US/60142896
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000052
; CURRENT APPLICATION NUMBER: US/60/142,896
; CURRENT FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 1584
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1491
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Drosophila
; US-60-142-896-1491

Query Match      29.0%; Score 113.5; DB 27; Length 62;
Best Local Similarity 36.2%; Pred. No. 0.0008;
Matches 21; Conservative 8; Mismatches 26; Indels 3; Gaps 2;

QY 11 KCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIK--ICAPGCVRLGLYLRNKKKVCVPRSKC 66
    |||||
DB 4 RCPANETFLAGGPDQTECA-TLGKPLVHRIRCPDGCYCNCNKGFAARNAAGTCIPLRRC 60

RESULT 15
US-60-145-138-701
; Sequence 701, Application US/60145138
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000062
; CURRENT APPLICATION NUMBER: US/60/145,138
; CURRENT FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 701

; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
US-10-174-151-4
; Sequence 4, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
US-10-174-151-4

Query Match      100.0%; Score 391; DB 25; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGLYLRNKKKVC 60
    |||||
DB 5 GFGGLGGRGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGLYLRNKKKVC 64

QY 61 VPRSKCG 67
    |||||
DB 65 VPRSKCG 71

RESULT 12
US-10-204-145-4
; Sequence 4, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
US-10-204-145-4

Query Match      100.0%; Score 391; DB 26; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGLYLRNKKKVC 60
    |||||
DB 5 GFGGLGGRGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGLYLRNKKKVC 64

QY 61 VPRSKCG 67
    |||||
DB 65 VPRSKCG 71

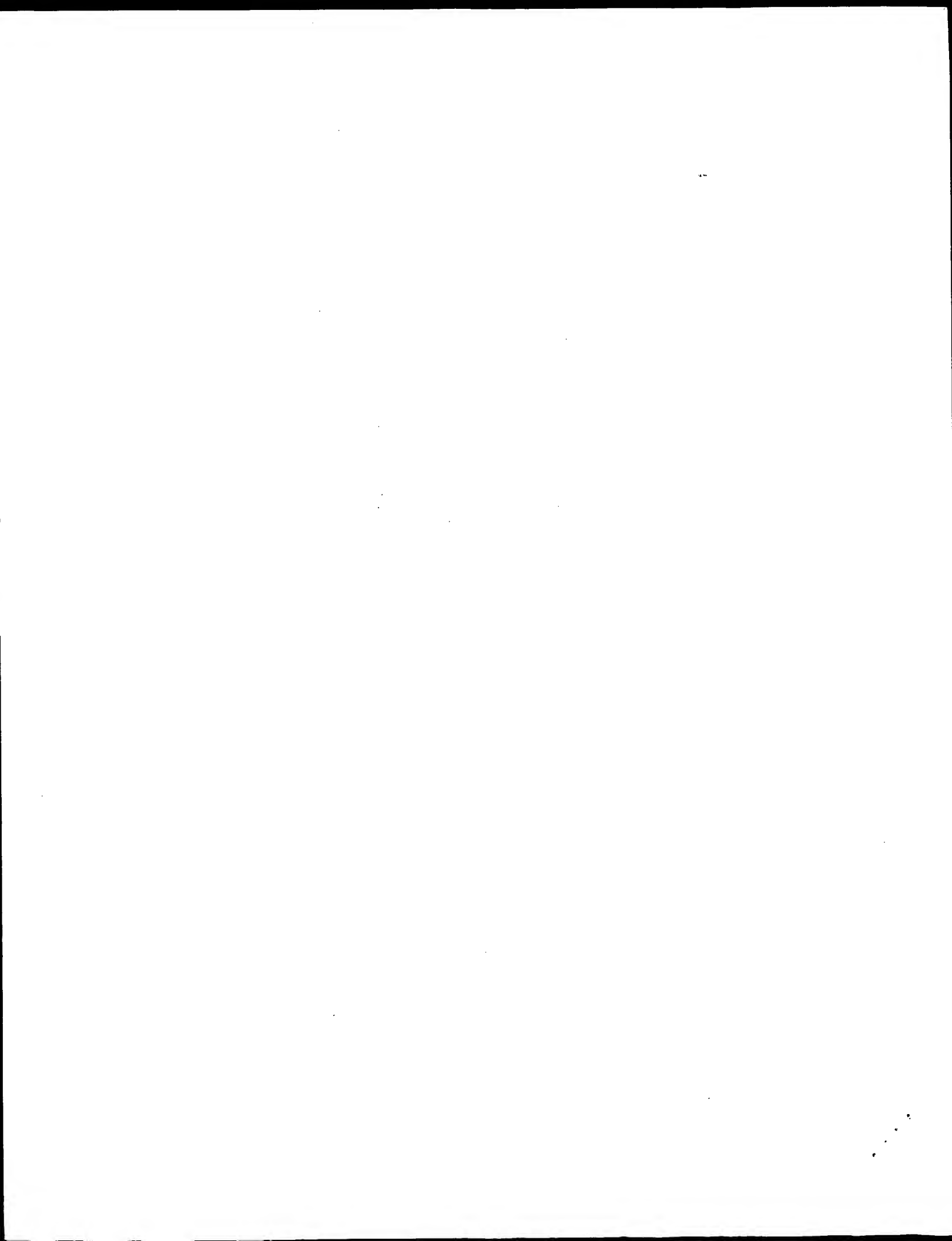
RESULT 13
PCT-US99-21077-2
```

; LENGTH: 62
; TYPE: PRT
; ORGANISM: Drosophila
US-60-145-138-701

Query Match 29.0%; Score 113.5; DB 27; Length 62;
Best Local Similarity 36.2%; Pred. No. 0.0008;
Matches 21; Conservative 8; Mismatches 26; Indels 3; Gaps 2;

Qy 11 KCPSEIFSRCDGRCPNVPKPLCIK--ICAPGCVCRGLYLRNKKKVCVPRSKC 66
:||:| | | | | : | | : | | | | : | | : |
Db 4 RCPANETFLACGPPDCOTECA-TLGKPCLVHRHICPDGCYCNKGFARNAAGTCIPLRRC 60

Search completed: February 26, 2003, 15:08:48
Job time : 142 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:04:32 ; Search time 23 seconds
(without alignments)
266.242 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGFGGLGGRGKCPNSIEFSR.....CRLGYLRNKKKVCVPSKCG 67

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 486122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New :

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	119	30.4	67	6	US-10-038-854-185	Sequence 185, App
2	113	28.9	5374	6	US-10-028-248A-75	Sequence 75, App
3	113	28.9	5374	6	US-10-107-782-75	Sequence 75, App
4	113	28.9	5376	6	US-10-028-248A-74	Sequence 74, App
5	113	28.9	5376	6	US-10-107-782-74	Sequence 74, App
6	109	27.9	2601	6	US-10-028-248A-76	Sequence 76, App
7	109	27.9	2601	6	US-10-107-782-76	Sequence 76, App
8	106.5	27.2	4360	5	US-09-724-676-66993	Sequence 66993, A
9	106.5	27.2	4360	5	US-09-724-676A-66993	Sequence 66993, A
10	98	25.1	2813	1	PCT-US02-04915-149	Sequence 149, App
11	97.5	24.9	111	1	PCT-US02-40655-379	Sequence 379, App
12	90	23.0	4219	6	US-10-085-198-2	Sequence 2, App
13	81	20.7	1082	6	US-10-264-237-2725	Sequence 2725, App
14	81	20.7	1458	1	PCT-US02-13209-25	Sequence 25, App
15	81	20.7	2212	6	US-10-028-248A-43	Sequence 43, App
16	81	20.7	2212	6	US-10-107-782-43	Sequence 43, App
17	81	20.7	2570	6	US-10-028-248A-42	Sequence 42, App
18	81	20.7	2570	6	US-10-107-782-42	Sequence 42, App
19	78	19.9	1568	6	US-10-263-929-117	Sequence 117, App
20	76.5	19.6	855	6	US-10-072-012-355	Sequence 355, App
21	76.5	19.6	855	6	US-10-072-012-413	Sequence 413, App
22	76.5	19.6	902	6	US-10-333-743-3	Sequence 3, App
23	75.5	19.3	5179	7	US-60-438-735-151	Sequence 151, App
24	73.5	18.8	855	6	US-10-072-012-356	Sequence 356, App
25	73.5	18.8	855	6	US-10-072-012-414	Sequence 414, App
26	73.5	18.8	855	6	US-10-072-012-417	Sequence 417, App

27	73.5	18.8	855	6	US-10-072-012-417	Sequence 417, App
28	73	18.7	319	6	US-10-017-161-1346	Sequence 1346, App
29	73	18.7	969	6	US-10-052-648A-34	Sequence 34, App
30	73	18.7	969	6	US-10-052-648A-35	Sequence 35, App
31	73	18.7	969	6	US-10-055-877-214	Sequence 214, App
32	72	18.4	366	5	US-09-963-693-103	Sequence 103, App
33	72	18.4	1367	5	US-09-751-708A-120	Sequence 120, App
34	72	18.4	1367	5	US-09-949-001-18	Sequence 18, App
35	72	18.4	1376	6	US-10-276-774-2170	Sequence 2170, App
36	72	18.4	1377	5	US-09-949-001-21	Sequence 21, App
37	72	18.4	1664	6	US-10-055-877-212	Sequence 212, App
38	71	18.2	1511	6	US-10-228-206-11	Sequence 11, App
39	71	18.2	1581	5	US-09-949-002-414	Sequence 414, App
40	71	18.2	1587	5	US-09-949-002-354	Sequence 354, App
41	71	18.2	1620	6	US-10-055-877-213	Sequence 213, App
42	70	17.9	1393	6	US-10-312-352-21	Sequence 21, App
43	70	17.9	1398	6	US-10-055-877-46	Sequence 46, App
44	70	17.9	1403	6	US-10-055-877-52	Sequence 52, App
45	70	17.9	1404	6	US-10-055-877-44	Sequence 44, App

ALIGNMENTS

RESULT 1

US-10-038-854-185
; Sequence 185, Application US/10038854
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca J
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29

```
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-185

Query Match      30.4%; Score 119; DB 6; Length 67;
Best Local Similarity 37.3%; Pred. No. 1.3e-05;
Matches 25; Conservative 7; Mismatches 23; Indels 12; Gaps 3;

QY 12 CPSNEIFSRCDGRCQRCFCPN-----VVP-----KPLCIKICAPGCVCRGLGRLNKK-KV 59
Db 1 CPANQYTCGGSCPSCSNPDGPLETTPPCGTSKVPSTCKECCVCQPGYVRNNDGDK 60
QY .60 CVPRSKC 66
Db 61 CVPRSEC 67

RESULT 2
US-10-028-248A-75
; Sequence 75, Application US/10028248A
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides and Methods of Use
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
```

```
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 5374
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-028-248A-75

Query Match      28.9%; Score 113; DB 6; Length 5374;
Best Local Similarity 37.5%; Pred. No. 0.0014;
Matches 24; Conservative 8; Mismatches 22; Indels 10; Gaps 3;

QY 11 KPSPNEIFSRCDGRCQRCFCPNVVPKPLC-----IKICAPGCVCRGLGRLNKKKVCVPR 63
Db 4741 KCPANSLYTHCLPTCLPSCSN--PDGRCEGTSHKAPSTCREGCVQPGYLLN-KDTCVHK 4797
QY 64 SKCG 67
Db 4798 NQCG 4801

RESULT 3
US-10-107-782-75
; Sequence 75, Application US/10107782
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREO
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 75
; LENGTH: 5374
```


; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66993
; LENGTH: 4360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66993

Query Match 27.2%; Score 106.5; DB 5; Length 4360;
Best Local Similarity 36.8%; Pred. No. 0.0057;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
QY 11 KCPNSNEIFSRCDGRCQRCFPCNVVPKPLCIKICAPGCVCRGLGYLRNKKKVCVPRSKCG 67
Db 1531 ECPNSHYELCADTCSLGCALSAPPQCDGCAEGCQCDGFLYN-GQACVPIQQCG 1586

RESULT 9
PCT-US02-676A-66993
; Sequence 66993, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66993
; LENGTH: 4360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66993

Query Match 27.2%; Score 106.5; DB 5; Length 4360;
Best Local Similarity 36.8%; Pred. No. 0.0057;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
QY 11 KCPNSNEIFSRCDGRCQRCFPCNVVPKPLCIKICAPGCVCRGLGYLRNKKKVCVPRSKCG 67
Db 1531 ECPNSHYELCADTCSLGCALSAPPQCDGCAEGCQCDGFLYN-GQACVPIQQCG 1586

RESULT 10
PCT-US02-04915-149
; Sequence 149, Application PC/TUS0204915
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynnne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200PC
; CURRENT APPLICATION NUMBER: PCT/US02/04915
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/285,475
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29

; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-04915-149

Query Match 25.1%; Score 98; DB 1; Length 2813;
Best Local Similarity 26.0%; Pred. No. 0.033;
Matches 19; Conservative 13; Mismatches 27; Indels 14; Gaps 2;

QY 8 GRC-----KCPNSNEIFSRCDGRCQRCFPCNVVPKPLCIKICAPGCVCRGLGYL 53
Db 635 GRGVRVAVREPGRCEINCPKGGVYLQCGTFCNLTCRSLSYPDCEACLEGCFCPPGLY 694
QY 54 RNKKKVCVPRSKC 66
Db 695 MDERGDCVPKAC 707

RESULT 11
PCT-US02-40655-379
; Sequence 379, Application PC/TUS0240655
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO PANCREATIC SPECIFIC GENES
; FILE REFERENCE: DEX-0376
; CURRENT APPLICATION NUMBER: PCT/US02/40655
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/342,768
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 543
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 379
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapien
PCT-US02-40655-379

Query Match 24.9%; Score 97.5; DB 1; Length 111;
Best Local Similarity 28.6%; Pred. No. 0.0035;
Matches 16; Conservative 13; Mismatches 26; Indels 1; Gaps 1;

QY 12 CPSNEIFSRCDGRCQRCFPCNVVPKPLCIKICAPGCVCRGLGYLRNKKKVCVPRSKC 66
Db 34 CPKGGVYLQCGTFCNLTCRSLSYPDCEACLEGCFCPPGLYMDERGDCVPKAC 89

RESULT 12
US-10-085-198-2
; Sequence 2, Application US/10085198
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20

```

RESULT 14
PCT-US02-13209-25
; Sequence 25, Application PC/TUS0213209
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H.
; APPLICANT: WEIGEL, JANET A.
; TITLE OF INVENTION: METHODS OF USING A HYALURONAN RECEPTOR
; FILE REFERENCE: 619758-8/JP/286.468
; CURRENT APPLICATION NUMBER: PCT/US02/13209

```

```

Query Match      20.7%; Score 81; DB 1; Length 1460;
Best Local Similarity 31.9%; pred. No. 1.3;
Matches 23; Conservative 9; Mismatches 22; Indels 18; Gaps 4;

Qy 5 GLGGRGKCPSENEIFS--RCDGRGCRFCFNNVVKPLCIKICAPGCVCR-----LGYLK 54
      ||||| | | : : || : : : ||| ||| |||
Db 922 GLGSGSCFCDEGWTGPRCEVQLEL-----QPVTCPPEAVCAVCRAGNSECSLGY-E 973

Qy 55 NKKKVCVPRSKC 66
      : || |
Db 974 GDGRVCTVADLC 985

```

Search completed: February 26, 2003, 15:09:19
Job time : 25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:03:23 : Search time 30 seconds
(without alignments)
65.711 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGFGGLGGRGKCPSENEIFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	391	100.0	92	4	US-09-394-630-2
2	107.5	27.5	2476	2	US-08-276-967-2
3	106.5	27.2	2594	4	US-08-718-388-7
4	106.5	27.2	5405	4	US-08-718-388-9
5	100	25.6	84	2	US-08-465-380-59
6	100	25.6	84	2	US-08-486-397-59
7	100	25.6	84	2	US-08-486-399-59
8	100	25.6	84	2	US-08-461-965-59
9	100	25.6	84	2	US-08-634-641-59
10	100	25.6	84	3	US-09-249-471-59
11	100	25.6	84	3	US-09-249-472-59
12	100	25.6	84	3	US-09-249-451-59
13	100	25.6	84	3	US-08-809-455-59
14	100	25.6	84	3	US-09-249-461-59
15	100	25.6	84	3	US-09-249-448-59
16	100	25.6	91	2	US-08-465-380-128
17	100	25.6	91	2	US-08-480-478-50
18	100	25.6	91	2	US-08-486-397-128
19	100	25.6	91	2	US-08-486-399-128
20	100	25.6	91	2	US-08-461-965-128
21	100	25.6	91	2	US-08-326-110A-50
22	100	25.6	91	2	US-08-634-641-128
23	100	25.6	91	3	US-09-249-471-128
24	100	25.6	91	3	US-09-249-472-128
25	100	25.6	91	3	US-09-249-451-128
26	100	25.6	91	3	US-08-809-455-128
27	100	25.6	91	3	US-09-249-461-128

28 100 25.6 91 3 US-09-249-448-128 Sequence 128, Appl
29 98 25.1 115 4 US-09-219-983A-4 Sequence 4, Appl
30 91 23.3 82 2 US-08-465-380-25 Sequence 25, Appl
31 91 23.3 82 2 US-08-465-380-49 Sequence 49, Appl
32 91 23.3 82 2 US-08-480-478-54 Sequence 54, Appl
33 91 23.3 82 2 US-08-486-397-25 Sequence 25, Appl
34 91 23.3 82 2 US-08-486-397-49 Sequence 49, Appl
35 91 23.3 82 2 US-08-486-399-25 Sequence 25, Appl
36 91 23.3 82 2 US-08-486-399-49 Sequence 49, Appl
37 91 23.3 82 2 US-08-461-965-25 Sequence 25, Appl
38 91 23.3 82 2 US-08-461-965-49 Sequence 49, Appl
39 91 23.3 82 2 US-08-326-110A-54 Sequence 54, Appl
40 91 23.3 82 2 US-08-634-641-25 Sequence 25, Appl
41 91 23.3 82 2 US-08-634-641-49 Sequence 49, Appl
42 91 23.3 82 3 US-09-249-471-25 Sequence 25, Appl
43 91 23.3 82 3 US-09-249-471-49 Sequence 49, Appl
44 91 23.3 82 3 US-09-249-472-25 Sequence 25, Appl
45 91 23.3 82 3 US-09-249-472-49 Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-09-394-630-2
; Sequence 2, Application US/09394630
; Patent No. 6395306
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Xuefeng
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200US
; CURRENT APPLICATION NUMBER: US/09/394,630
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 60/100,172
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Apis mellifera
US-09-394-630-2

Query Match 100.0%; Score 391; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.2e-33;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGFGGLGGRGKCPSENEIFSRCDGRCPNNVVKPLIKICAPGCVRLGLRNKKVC 60
Db 26 GGFGGLGGRGKCPSENEIFSRCDGRCPNNVVKPLIKICAPGCVRLGLRNKKVC 85
Qy 61 VPRSKG 67
Db 86 VPRSKG 92

RESULT 2
US-08-276-967-2
; Sequence 2, Application US/08276967
; Patent No. 5851817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; TITLE OF INVENTION: Sperm
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston

```
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276.967
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UTSD:418KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2476 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-967-2

Query Match 27.5%; Score 107.5; DB 2; Length 2476;
Best Local Similarity 33.3%; Pred. No. 0.0038;
Matches 22; Conservative 9; Mismatches 18; Indels 17; Gaps 3;

QY 11 KCPNSEIFSRCDGR-----QRCFNVVVKPLCIKICAPGCVRLGYLRNKKKVCV 61
DB 1455 KCPGSSYSTCANPCATCLSLNNPSYCPSTLP-----CAEGCEQKQHILSGTS-CV 1506

QY 62 PRSKCG 67
DB 1507 PLSQCG 1512

RESULT 3
US-08-718-388-7
; Sequence 7, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:
; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING IgG FC REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5405 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-718-388-9

Query Match 27.2%; Score 106.5; DB 4; Length 5405;
Best Local Similarity 36.8%; Pred. No. 0.0099;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;

QY 11 KCPNSEIFSRCDGRQCFPCNVVVKPLCIKICAPGCVRLGYLRNKKKVCVPRSKCG 67
DB 1531 ECPNPNHYELCADTCSLGCALSAPPQCQDGCAGCQCDSGFLYN-GQACVPIQCG 1586

RESULT 5
US-08-465-380-59
; Sequence 59, Application US/08465380
; Patent No. 5863894
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-718-388-7

Query Match 27.2%; Score 106.5; DB 4; Length 2594;
Best Local Similarity 36.8%; Pred. No. 0.005;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;

QY 11 KCPNSEIFSRCDGRQCFPCNVVVKPLCIKICAPGCVRLGYLRNKKKVCVPRSKCG 67
DB 1531 ECPNPNHYELCADTCSLGCALSAPPQCQDGCAGCQCDSGFLYN-GQACVPIQCG 1586

RESULT 4
US-08-718-388-9
; Sequence 9, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:
; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING IgG FC REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5405 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-718-388-9

Query Match 27.2%; Score 106.5; DB 4; Length 5405;
Best Local Similarity 36.8%; Pred. No. 0.0099;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;

QY 11 KCPNSEIFSRCDGRQCFPCNVVVKPLCIKICAPGCVRLGYLRNKKKVCVPRSKCG 67
DB 1531 ECPNPNHYELCADTCSLGCALSAPPQCQDGCAGCQCDSGFLYN-GQACVPIQCG 1586

RESULT 5
US-08-465-380-59
; Sequence 59, Application US/08465380
; Patent No. 5863894
; TELECOMMUNICATION INFORMATION:
```


GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVKPLCIKICAPCCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFFRNKD 61

QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70

RESULT 6
US-08-486-397-59
Sequence 59, Application US/08486397
Patent No. 5866542
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVKPLCIKICAPCCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFFRNKD 61

QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70

RESULT 7
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVKPLCIKICAPCCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFFRNKD 61

QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70

RESULT 7
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVKPLCIKICAPCCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFFRNKD 61

QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70

RESULT 7
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVKPLCIKICAPCCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFFRNKD 61

QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70

RESULT 7
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59

```

;
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,399
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-486-399-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVPKPLCIKICAPGCVCRGLYLRNKK 57
Db 5 QCGENEKYDCSGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFYRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 8
US-08-461-965-59
; Sequence 59, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/461,965
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 210/243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-461-965-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVPKPLCIKICAPGCVCRGLYLRNKK 57
Db 5 QCGENEKYDCSGKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFYRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 9
US-08-634-641-59
; Sequence 59, Application US/08634641
; Patent No. 5955294
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George P. Vlasuk
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Mensens, Joris Hilda Llieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,641
; FILING DATE: April 19, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231

```

; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 219/136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-634-641-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KPFSNEIFSRCDGR-CORFC-----PNVVPKPLCIKICAPGCVCRGLGYNKK 57
Db 5 QCGENEKYDCSGSKEDCKKCYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFGYRNKD 61
QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 10
US-09-249-471-59
; Sequence 59, Application US/09249471
; Patent No. 6040441
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Llieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,471
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-471-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KPFSNEIFSRCDGR-CORFC-----PNVVPKPLCIKICAPGCVCRGLGYNKK 57
Db 5 QCGENEKYDCSGSKEDCKKCYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFGYRNKD 61
QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 11
US-09-249-472-59
; Sequence 59, Application US/09249472
; Patent No. 6046318
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Llieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles

APPLICANT: Messens, Joris Hilda Lieven
 APPLICANT: Lauwereys, Marc Josef
 APPLICANT: Laroche, Yves Rene
 APPLICANT: Jespers, Laurent Stephane
 APPLICANT: Gansemans, Yannick Georges Jozef
 APPLICANT: Moyle, Matthew
 APPLICANT: Bergum, Peter W.
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
 TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
 NUMBER OF SEQUENCES: 356
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Suite 4700
 STATE: Los Angeles
 COUNTRY: U.S.A.
 ZIP: 90071

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/809,455
 FILING DATE: April 17, 1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/13231
 FILING DATE: October 17, 1995
 APPLICATION NUMBER: 08/486,399
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/486,397
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/465,380
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/461,965
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 216/270
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 TYPE: amino acid
 LENGTH: 84 amino acids
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Ancylostoma caninum
 US-08-809-455-59

Query Match 25.6%; Score 100; DB 3; Length 84;
 Best Local Similarity 30.4%; Pred. No. 0.00097;
 Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPNSIFSRCDGR-CQRFC-----PNVVPKPLCIKICAPGCVRLGLYLRNKK 57
 Db 5 QCGENEKYSKSGKEDKCKKYDGVVEEDDEPNV---PCLVRVCHQDCVCEEGFYRNKD 61

QY 58 KVCVPRSK 66

Db 62 DKCVSAEDC 70

RESULT 14

US-09-249-461-59
 Sequence 59, Application US/09249461
 Patent No. 6096877
 GENERAL INFORMATION:
 APPLICANT: Vlasuk, George Phillip
 APPLICANT: Staussens, Patrick Eric Hugo
 APPLICANT: Messens, Joris Hilda Lieven
 APPLICANT: Lauwereys, Marc Josef
 APPLICANT: Laroche, Yves Rene
 APPLICANT: Jespers, Laurent Stephane
 APPLICANT: Gansemans, Yannick Georges Jozef
 APPLICANT: Moyle, Matthew
 APPLICANT: Bergum, Peter W.
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
 TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
 NUMBER OF SEQUENCES: 356
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Suite 4700
 STATE: Los Angeles
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/249,461
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/809,455
 FILING DATE: April 17, 1997
 APPLICATION NUMBER: PCT/US95/13231
 FILING DATE: October 17, 1995
 APPLICATION NUMBER: 08/486,399
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/486,397
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/465,380
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/461,965
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 216/270
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 84 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Ancylostoma caninum
 US-09-249-461-59

Query Match 25.6%; Score 100; DB 3; Length 84;
 Best Local Similarity 30.4%; Pred. No. 0.00097;
 Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPNSIFSRCDGR-CQRFC-----PNVVPKPLCIKICAPGCVRLGLYLRNKK 57
 Db 5 QCGENEKYSKSGKEDKCKKYDGVVEEDDEPNV---PCLVRVCHQDCVCEEGFYRNKD 61

Wed Feb 26 15:12:38 2003

US-09-249-448-59

Query Match
Best Local Similarity 25.6%; Score 100; DB 3; Length 84;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSTNEIFSRCDGR-CQRF-----PNNVYPKPLCIKICAPGCVCLGYLRNKK 57
Db 5 QCGENEKYDSCGSKCEDKCKKYDGVVEEDDEEPNV---PCLVRVCHQDCVCEGFGYRNKD 61
QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

Search completed: February 26, 2003, 15:06:20
Job time : 32 secs

US-09-249-448-59

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 15
US-09-249-448-59
: Sequence 59, Application US/09249448
: Patent No. 6121435
: GENERAL INFORMATION:
: APPLICANT: Vlasuk, George Phillip
: APPLICANT: Stanssens, Patrick Eric Hugo
: APPLICANT: Messens, Joris Hilda lleven
: APPLICANT: Lauwereys, Marc Josef
: APPLICANT: Laroche, Yves Rene
: APPLICANT: Jespers, Laurent Stephane
: APPLICANT: Gansemans, Yannick Georges Jozef
: APPLICANT: Moyle, Matthew
: APPLICANT: Bergum, Peter W.
: TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
: TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
: TITLE OF INVENTION: PROTEIN
: NUMBER OF SEQUENCES: 356
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/249,448
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/809,455
: FILING DATE: April 17, 1997
: APPLICATION NUMBER: PCT/US95/13231
: FILING DATE: October 17, 1995
: APPLICATION NUMBER: 08/486,399
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/486,397
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/465,380
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/461,965
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/326,110
: FILING DATE: October 18, 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: BIGGS, SUZANNE L.
: REGISTRATION NUMBER: 30,158
: REFERENCE/DOCKET NUMBER: 216/270
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: INFORMATION FOR SEQ ID NO: 59:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 84 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Ancylostoma caninum

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:09:22 : Search time 17 seconds
(without alignments)
378.883 million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGGLGRGKCPSEIFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 19399

Minimum DB seq length: 0
Maximum DB seq length: 67

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	80.5	20.6	63	2 S07127	chymotrypsin/elast
2	77.5	19.8	63	2 S08572	chymotrypsin/elast
3	70.5	18.0	62	2 S35098	trypsin inhibitor
4	69	17.6	60	2 S31723	metallothionein -
5	69	17.6	60	2 B27490	metallothionein B
6	68	17.4	60	2 S30567	metallothionein -
7	67	17.1	62	2 A35640	metallothionein 4
8	65	16.6	62	2 B33640	metallothionein IV
9	62	15.9	61	2 A37425	metallothionein 2
10	62	15.9	62	2 S43335	metallothionein-2c
11	62	15.9	62	2 S43336	metallothionein-2a
12	61.5	15.7	61	2 A33889	metallothionein 1
13	60.5	15.5	61	2 S08572	metallothionein 1A
14	60.5	15.5	61	2 S08572	metallothionein A
15	60	15.3	52	2 S65712	metallothionein 1B
16	59.5	15.2	61	1 SMHUIB	metallothionein 1R
17	59.5	15.2	61	2 S69277	metallothionein 1R
18	59.5	15.2	63	2 S08190	metallothionein 1
19	59.5	15.2	63	2 A34905	metallothionein 1
20	59	15.1	60	2 JC2420	metallothionein -
21	59	15.1	61	1 SMBO2	metallothionein II
22	59	15.1	61	2 S0808	metallothionein Ia
23	59	15.1	61	2 S0809	metallothionein Ib
24	59	15.1	65	2 A38739	metallothionein -
25	58.5	15.0	61	1 S08572	metallothionein I
26	58.5	15.0	61	2 S47651	metallothionein 1H
27	58.5	15.0	61	2 I54574	metallothionein-1
28	58	14.8	61	2 S0810	metallothionein Ic
29	57.5	14.7	60	1 SMHUIA	metallothionein 1A

30	57.5	14.7	60	2 S38335	metallothionein -
31	57.5	14.7	61	1 SMHUIE	metallothionein 1E
32	57.5	14.7	61	1 SMHOB	metallothionein 1B
33	57.5	14.7	61	2 S47652	metallothionein 1X
34	56.5	14.5	61	1 SMHUI2	metallothionein 2
35	56.5	14.5	61	1 SMHUI2	metallothionein 2
36	56.5	14.5	61	1 SMHUI2	metallothionein 2
37	56	14.3	64	2 A33825	metallothionein 1F
38	55.5	14.2	61	1 SMRT1	metallothionein I
39	55	14.1	64	2 A25775	metallothionein A
40	54.5	13.9	48	2 G01478	thrombospondin-p50
41	54.5	13.9	61	1 SMHUI2C	metallothionein II
42	54.5	13.9	61	2 S0811	metallothionein II
43	54.5	13.9	61	2 B23889	metallothionein 2
44	54.5	13.9	61	2 I48173	metallothionein II
45	54.5	13.9	61	2 I48116	metallothionein II

ALIGNMENTS

RESULT 1

S07127

chymotrypsin/elastase inhibitor - common roundworm

C:Species: Ascaris lumbricoides (common roundworm)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997

C:Accession: S07127

R:Babin, D.R.; Peanasky, R.J.; Goos, S.M.

A:Title: The inhibitors of chymotrypsin/elastase from Ascaris lumbricoides: the pr

A:Reference number: S07127; MUID:84255715; PMID:6564898

A:Accession: S07127

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-63 <BAB>

C:Superfamily: roundworm trypsin inhibitor

Query Match 20.6%; Score 80.5; DB 2; Length 63;

Best Local Similarity 30.3%; Pred. No. 0.23;

Matches 20; Conservative 10; Mismatches 23; Indels 13; Gaps 4;

QY 8 GRGKCPSEIFSRCDG---RC-----QRFCNVVVPKPLCIKICAPGCVRLGVRNKKKVC 60

DB 1 QGSCGPNVWTECTGCKGKCPDENTPCPLMCRPSC--ECSPG-----RGMRTNDGKC 54

QY 61 VPRSKC 66

DB 55 IPASQC 60

RESULT 2

S08572

chymotrypsin/elastase inhibitor - common roundworm

C:Species: Ascaris lumbricoides (common roundworm)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997

C:Accession: S08572

R:Babin, D.R.; Peanasky, R.J.; Goos, S.M.

A:Title: The inhibitors of chymotrypsin/elastase from Ascaris lumbricoides: the pr

A:Reference number: S07127; MUID:84255715; PMID:6564898

A:Accession: S08572

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-63 <BAB>

C:Superfamily: roundworm trypsin inhibitor

Query Match 19.8%; Score 77.5; DB 2; Length 63;

Best Local Similarity 31.7%; Pred. No. 0.45;

Matches 19; Conservative 6; Mismatches 32; Indels 3; Gaps 2;

QY 9 RKGKCPSEIFSRCDGRCQRFCNVVVPKPLCIKICAPGCVCR--LGVRNKKKVCVPRSKC 66

DB 1 RKPCGKNVWTECTG--CELKCGQDENTPCALMCRPPSCSPGRMRTHDGKCVPVSEC 59

C; Superfamily: metalloionin
Query Match 15.7%; Score 61.5; DB 2; Length 6
Best Local Similarity 36.7%; Pred. No. 16;
Matches 18; Conservative 5; Mismatches 19; Indels

Db	<pre> 11 : : : : : : : : : : 10 GGSCTCAGCKCKA-----CRPS-CKKSCCSPVG-CAK-CAQGCVC 51 </pre>	<p>C:Superfamily: metallothionein C:Keywords: metal binding</p>
RESULT 13	<pre> SMHUIA metallothionein 1A - human C:Species: Homo sapiens (man) C>Date: 24-Jun-1987 #sequence_revision 09-Sep-1994 #text_change 22-Jun-1999 C:Accession: A24502 R:Richards, R.I.; Heguy, A.; Karin, M. Cell 37, 263-272, 1984 </pre>	<p>Query Match 15.5%; Score 60.5; DB 1; Length 61; Best Local Similarity 35.0%; Pred. No. 20; Matches 14; Conservative 7; Mismatches 14; Indels 5; Gaps 4;</p>
QY	<pre> 10 GKPCSNFISRCDCGRFCPNVVPKPLCIKICAPGCVCR 49 : : : : : : : </pre>	<p>QY 7 GGRGKCPSNEIFSRCDGRCQRFPCPNVVPKPLCIKICAPGCVCR 49 Db 8 GGSCTCSSCGCKNCKKSCCCPVGCSK-----CAQGCVC 44</p>
Db	<pre> 17 GSCKCKE--CKNS-CKKSCCSPMS-CAK-CAQGCICK 51 </pre>	<p>Search completed: February 26, 2003, 15:13:00 Job time : 17 secs</p>

FT METAL 35 35 CLUSTER A.
 FT METAL 36 36 CLUSTER A.
 FT METAL 40 40 CLUSTER A.
 FT METAL 43 43 CLUSTER A.
 FT METAL 47 47 CLUSTER A.
 FT METAL 49 49 CLUSTER A.
 FT METAL 54 54 CLUSTER A.
 FT METAL 58 58 CLUSTER A.
 FT METAL 59 59 CLUSTER A.
 SQ SEQUENCE 60 AA; 5926 MW; E862A5EC16D7348A CRC64;

Query Match 18.7%; Score 73; DB 1; Length 60;
 Best Local Similarity 35.0%; Pred. No. 0.17;
 Matches 21; Conservative 3; Mismatches 22; Indels 14; Gaps 3;

QY 1 GGTGGLGGRGRCQRCPCPNVVKPLGICAPCCVRLGLRNKKKVC 60
 DB 9 GGTGCGGSCVC-TNCSTTCKKSCCPGCGCPA-----CAGGVC-----KGKTC 54

RESULT 5
 ITRI_ASCSU STANDARD; PRT; 62 AA.
 AC P19398;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Trypsin inhibitor (ATI).
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
 OC Ascarididae; Ascaris.
 OX NCBI_TaxID=6253;
 RN [1]
 RP SEQUENCE.
 RA Peanasky R.J., Martzen M.R., Homandberg G.A., Cash J.M., Babin D.R.,
 RA Litwiller R.D.;
 RL (In) McInnis A.J. (eds.);
 RL Paradigms for eradicating helminthic parasites, pp.349-366,
 RL Alan R. Liss, New York (1987).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=90212567; PubMed=2322539;
 RA Gronenborn A.M., Nilges M., Peanasky R.J., Clore G.M.;
 RT "Sequential resonance assignment and secondary structure
 RT determination of the Ascaris trypsin inhibitor, a member of a novel
 RT class of proteinase inhibitors.";
 RL Biochemistry 29:183-189(1990).
 CC -!- SIMILARITY: BELONGS TO THE ASCARIS TRYPSIN INHIBITOR FAMILY.
 DR PIR; S35098.
 DR PDB; 1ATA; 31-AUG-94.
 DR PDB; 1ATB; 31-AUG-94.
 DR PDB; 1ATD; 31-AUG-94.
 DR PDB; 1ATE; 31-AUG-94.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR Pfam; PF01826; TIL; 1.
 KW Serine protease inhibitor; 3D-structure.
 FT DISULFID 5 38
 FT DISULFID 15 33
 FT DISULFID 18 29
 FT DISULFID 22 60
 FT DISULFID 40 54
 SQ SEQUENCE 62 AA; 6798 MW; 6438D25D2F76D3E1 CRC64;

Query Match 18.0%; Score 70.5; DB 1; Length 62;
 Best Local Similarity 29.5%; Pred. No. 0.33;
 Matches 18; Conservative 12; Mismatches 22; Indels 9; Gaps 5;

QY 11 KCFPS-NEIFSRCDGRCQRCPCPNVVKPLGICAP---GCVCRLGLRNKKKVCVPRSK 65
 DB 4 KCTRPNQWTKCGG-CEGTCAQRIIVP---CTRCCKPRCECIASAGFVRDAQNCIKFED 59
 QY 66 C 66

Db 60 C 60

RESULT 6
 MTB_ONCMY STANDARD; PRT; 60 AA.
 AC P09862;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein B (MT-B).
 GN MTB.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri), and
 OS Salvelinus alpinus (Arctic char).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022, 8036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.mykiss;
 RX MEDLINE=88111026; PubMed=2448099;
 RA Bonham K., Zafarullah M., Gedamu L.;
 RT "The rainbow trout metallothioneins: molecular cloning and
 RT characterization of two distinct cDNA sequences.";
 RL DNA 6:519-528(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.mykiss;
 RX MEDLINE=89039876; PubMed=3185557;
 RA Zafarullah M., Bonham K., Gedamu L.;
 RT "Structure of the rainbow trout metallothionein B gene and
 RT characterization of its metal-responsive region.";
 RL Mol. Cell. Biol. 8:4469-4476(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.mykiss; TISSUE=Liver;
 RX MEDLINE=91316146; PubMed=1859844;
 RA Kille P., Stephens P.E., Kay J.;
 RT "Elucidation of cDNA sequences for metallothioneins from rainbow
 RT trout, stone loach and pike liver using the polymerase chain
 RT reaction.";
 RL Biochim. Biophys. Acta 1089:407-410(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.alpinus; TISSUE=Liver;
 RA Gerpe M., Kling P., Olsson P.E.;
 RT "Metallothionein cDNA sequences and gene expression in arctic char
 RT (Salvelinus alpinus) following metal and PCB exposure.";
 RL Mar. Environ. Res. 46:551-554(1998).
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M18104; AAA49566.1; -
 CC EMBL; M22487; AAA49567.1; -
 CC EMBL; X59394; CAA42037.1; -
 CC EMBL; AF013801; AAB66343.1; -
 CC PIR; A30818; A30818.
 CC PIR; B27490; B27490.

```

DR PIR: S16997; S16997.
DR PIR: S15509; S15509.
DR HSP: P28184; L1J19.
DR InterPro; IPR003019; Metallthion.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PS00860; MTVERTEBRATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster.
FT DOMAIN 1 28
FT METAL 29 60
FT METAL 4 4
FT METAL 6 6
FT METAL 12 12
FT METAL 14 14
FT METAL 18 18
FT METAL 20 20
FT METAL 23 23
FT METAL 25 25
FT METAL 28 28
FT METAL 32 32
FT METAL 33 33
FT METAL 35 35
FT METAL 36 36
FT METAL 40 40
FT METAL 43 43
FT METAL 47 47
FT METAL 49 49
FT METAL 54 54
FT METAL 58 58
FT METAL 59 59
SQ SEQUENCE 60 AA; 6033 MW; 9EA1E70FBES9B4EE CRC64;

Query Match 17.6%; Score 69; DB 1; Length 60;
Best Local Similarity 37.0%; Pred. No. 0.46; Indels 14; Gaps 3;
Matches 20; Conservative 2; Mismatches 18;

QY 7 GGRGKCPSEIFSRGDCRCQFCFNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
DB 15 GGSCKC-SNCACTSKCKSCCPCCPSCKS-----CASGCV-----KGKTC 54

RESULT 7
MT_ESOLU
ID MT_ESOLU STANDARD; PRT; 60 AA.
AC P25127;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT).
GN MT.
OS Exo lucius (Northern pike).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Esociformes;
OC Esocidae; Esoc.
OX NCBI_TaxID=8010;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91316146; PubMed=1859844;
RA Kille P., Stephens P.E., Kay J.;
RT "Elucidation of cDNA sequences for metallothioneins from rainbow
RT trout, stone loach and pike liver using the polymerase chain
RT reaction.";
RL Biochim. Biophys. Acta 1089:407-410(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94032489; PubMed=8218416;
RA Kille P., Kay J., Sweeney G.E.;
RT "Analysis of regulatory elements flanking metallothionein genes in
RT Cd-tolerant fish (pike and stone loach).";
RL Biochim. Biophys. Acta 1216:55-64(1993).
KW FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
-!- RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
-!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
-!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
-!- This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X59392; CAA42035.1; -
EMBL; X70042; CAA49636.1; -
PIR; S17175; S17175.
PIR; S15503; S15503.
PIR; S38334; S38334.
PIR; S31723; S31723.
HSP; P02795; 1MHU.
InterPro; IPR003019; Metallthion.
InterPro; IPR000006; Metallthion_vert.
Pfam; PF00131; metalthio; 1.
PRINTS; PR00860; MTVERTEBRATE.
PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster.
FT DOMAIN 1 28
FT METAL 29 60
FT METAL 4 4
FT METAL 6 6
FT METAL 12 12
FT METAL 14 14
FT METAL 18 18
FT METAL 20 20
FT METAL 23 23
FT METAL 25 25
FT METAL 28 28
FT METAL 32 32
FT METAL 33 33
FT METAL 35 35
FT METAL 36 36
FT METAL 40 40
FT METAL 43 43
FT METAL 47 47
FT METAL 49 49
FT METAL 54 54
FT METAL 58 58
FT METAL 59 59
SQ SEQUENCE 60 AA; 5979 MW; 9EA1E43F95F8D97E CRC64;

Query Match 17.6%; Score 69; DB 1; Length 60;
Best Local Similarity 35.2%; Pred. No. 0.46; Indels 14; Gaps 3;
Matches 19; Conservative 3; Mismatches 18;

QY 7 GGRGKCPSEIFSRGDCRCQFCFNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
DB 15 GGSCKC-SNCACTSKCKSCCPCCPSCKS-----CASGCV-----KGKTC 54

RESULT 8
MT_PLEPL
ID MT_PLEPL STANDARD; PRT; 60 AA.
AC P07216;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT).
GN MT.
OS Pleuronectes platessa (Plaice).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


Db 11 GGICMGDNCKCTTCNCKTKRKSQCCP-PCGAK-CARGCICKGG---SDKSCSCP 62

```

RESULT 10
MT4_HUMAN          STANDARD;          PRT;          62 AA.
ID MT4_HUMAN
AC P47944;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-IV (MT-IV).
GN MT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94271779; PubMed=8003488;
RA Qualif C.J., Findley S.D., Erickson J.C., Froelick G.J.,
RA Kelly E.J., Zambrowicz B.P., Palminter R.D.;
RT "Induction of a new metallothionein isoform (MT-IV) occurs during
RT differentiation of stratified squamous epithelia.";
RL Biochemistry 33:7250-7259(1994).
CC -!- FUNCTION: SEEMS TO BIND ZINC AND COPPER. COULD PLAY A SPECIAL ROLE
CC IN REGULATING ZINC METABOLISM DURING THE DIFFERENTIATION OF
CC STRATIFIED EPITHELIA.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U07807; AAA20232.1; -
DR HSSP; P18055; 1MRB.
DR Genew; HGNC:18705; MT4.
DR MIM; 606206; -
DR InterPro; IPR003019; Metallthion.
DR InterPro; IPR000006; Metallthion_vert.
DR Pfam; PF00131; metalthio: 1.
DR PRINTS; PR00860; MTVERTBERATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Zinc; Copper.
FT METAL 6
FT METAL 8
FT METAL 14
FT METAL 16
FT METAL 20
FT METAL 22
FT METAL 25
FT METAL 27
FT METAL 30
FT METAL 34
FT METAL 35
FT METAL 37
FT METAL 38
FT METAL 42
FT METAL 45
FT METAL 49
FT METAL 51
FT METAL 58
FT METAL 60
FT METAL 61
SQ SEQUENCE 62 AA; 6419 MW; 36157CBA17BF28CC CRC64;

Query Match 17.1%; Score 67; DB 1; Length 62;
Best Local Similarity 35.1%; Pred. No. 0.76; Indels 4;
Matches 20; Conservative 8; Mismatches 23; Gaps 4;

7 GGRKGKPSNEIFSRCDGR-CQRFPCNVVPKPLCIKICAPGCVCRGLYLRNKKKVCVP 62

RESULT 11
MT_BALMY
ID MT_BALMY          STANDARD;          PRT;          61 AA.
AC O18842;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT).
OS Balaena mysticetus (Bowhead whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti; Balaenidae;
OX Balaena.
OX NCBI_TaxID=27602;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Kayser J., O'Hara T., Goodwin T., Linnehan R., Hammond T.;
RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RL Submitters HAVE A HIGH CONTENT OF CYSTEINE
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTINES.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022117; AAB72006.1; -
DR HSSP; P02795; 2MHU.
DR InterPro; IPR003019; Metallthion.
DR InterPro; IPR000006; Metallthion_vert.
DR Pfam; PF00131; metalthio: 1.
DR PRINTS; PR00860; MTVERTBERATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Acetylation.
FT DOMAIN 1
FT METAL 29
FT METAL 30
FT METAL 61
FT METAL 5
FT METAL 7
FT METAL 13
FT METAL 15
FT METAL 19
FT METAL 21
FT METAL 24
FT METAL 26
FT METAL 29
FT METAL 33
FT METAL 34
FT METAL 36
FT METAL 37
FT METAL 41
FT METAL 44
FT METAL 48
FT METAL 50
FT METAL 57
FT METAL 59
FT METAL 60
FT METAL 61
SQ SEQUENCE 61 AA; 6025 MW; 4D0A7C5E1D234A43 CRC64;

```

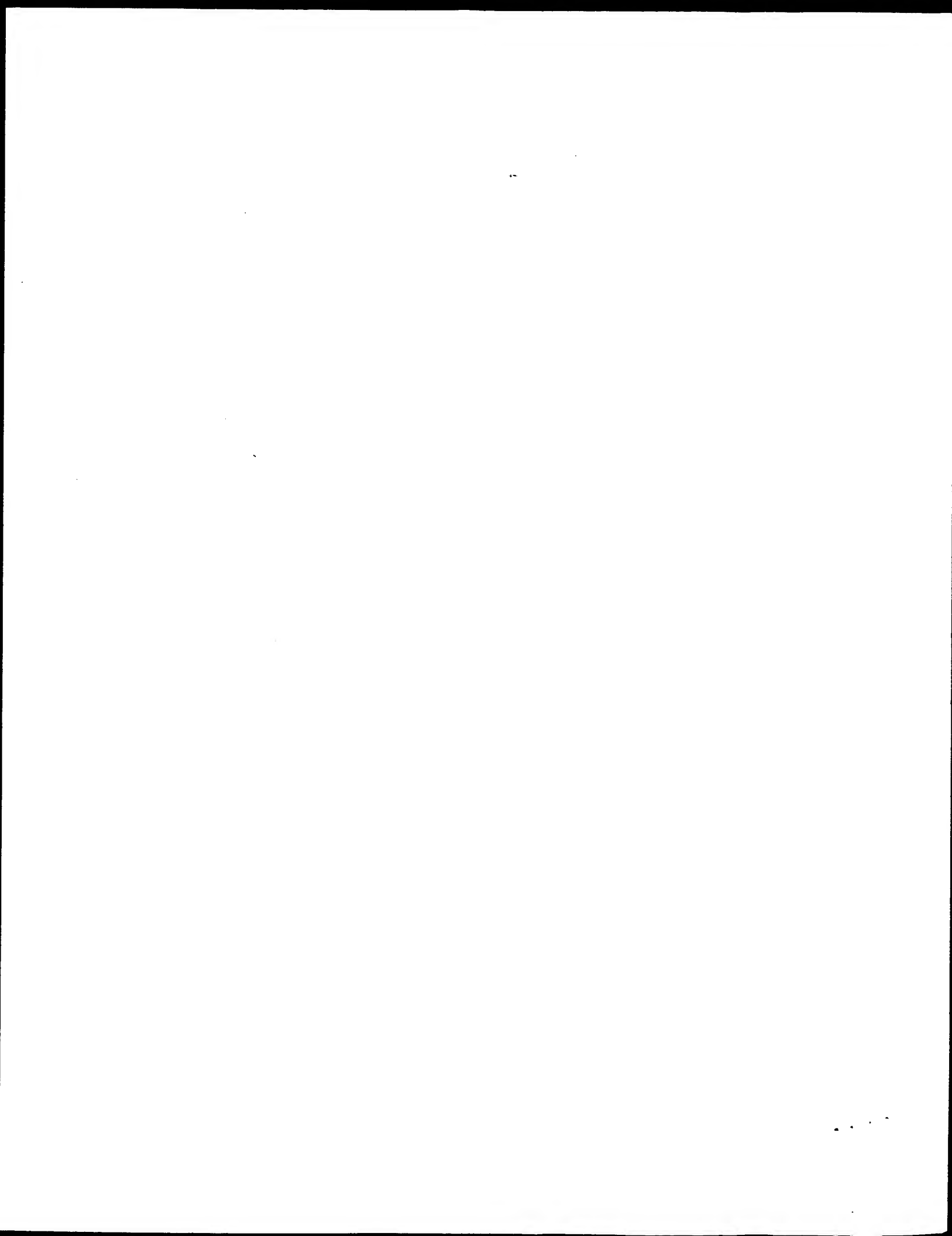


```
DR PRINTS: PR00860; MTVERTERATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster.
FT DOMAIN 1 28 BETA.
FT DOMAIN 29 60 ALPHA.
FT METAL 3 3 CLUSTER B.
FT METAL 5 5 CLUSTER B.
FT METAL 11 11 CLUSTER B.
FT METAL 13 13 CLUSTER B.
FT METAL 17 17 CLUSTER B.
FT METAL 19 19 CLUSTER B.
FT METAL 22 22 CLUSTER B.
FT METAL 24 24 CLUSTER B.
FT METAL 27 27 CLUSTER B.
FT METAL 31 31 CLUSTER A.
FT METAL 32 32 CLUSTER A.
FT METAL 34 34 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 39 39 CLUSTER A.
FT METAL 42 42 CLUSTER A.
FT METAL 46 46 CLUSTER A.
FT METAL 48 48 CLUSTER A.
FT METAL 56 56 CLUSTER A.
FT METAL 58 58 CLUSTER A.
FT METAL 59 59 CLUSTER A.
SQ SEQUENCE 60 AA; 5979 MW; AB05F49153953279 CRC64;

Query Match 16.2%; Score 63.5; DB 1; Length 60;
Best Local Similarity 33.3%; Pred. No. 1.7;
Matches 17; Conservative 4; Mismatches 23; Indels 7; Gaps 2;

Oy 1 GGFGGLGGRGKPSNEIFSRCDGRCPNVPKPLKIKICAPGCVCLRG 51
Db 8 GGSUSCAGSKCKNCKTS-CKKSCCCPSECEK-----CGGCGVCKGG 51

Search completed: February 26, 2003, 15:10:59
Job time : 14 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:08:52 ; Search time 91 Seconds
(without alignments)
151.705 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGLGGGKCPSEIFSR.....CRLGLRNKKKVCVPRKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 56343

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.5	22.6	62	5 077419	077419 ascaris suu
2	63.5	16.2	61	4 08WVB5	08wvb5 homo sapien
3	63	16.1	51	13 09S78	09s78 gallus gall
4	62.5	16.0	40	13 098TP9	098tp9 platichthys
5	62.5	16.0	44	11 099KF9	099kf9 mus musculus
6	60	15.3	61	4 08TDN3	08tdn3 homo sapien
7	59.5	15.2	63	4 09BSK9	09bsk9 homo sapien
8	59	15.1	48	13 098TC0	098tc0 seriola qui
9	58.5	15.0	58	4 09BQ2	09bq2 homo sapien
10	58.5	15.0	61	4 08TE66	08te66 homo sapien
11	58	14.8	60	13 08UVY1	08uvy1 lithognathu
12	57.5	14.7	49	13 098S15	098s15 leuciscus c
13	54.5	13.9	48	4 012928	012928 homo sapien
14	54.5	13.9	61	4 08TDC4	08tdc4 homo sapien
15	53	13.6	39	13 09PVG7	09pv7 xiphophorus
16	53	13.6	59	16 08XPB4	08xpb4 clostridium

17	53	13.6	66	4 08TCS1	08tcs1 homo sapien
18	51.5	13.2	59	5 09M9H1	09m9h1 ruditapes d
19	51.5	13.2	62	10 065918	065918 picea maria
20	51	13.0	48	16 08YUR2	08yur2 anabaena sp
21	51	13.0	65	17 08TNN8	08tnn8 methanosarc
22	50.5	12.9	58	5 095P38	095p38 homarus ame
23	50	12.8	58	5 095U91	095u91 scylla serr
24	50	12.8	67	16 08ZK16	08zki6 salmonella
25	49.5	12.7	55	17 097BH4	097bh4 thermoplasm
26	49.5	12.7	58	5 090623	090623 pacifastacu
27	49	12.5	58	5 095093	095u93 portunus pe
28	49	12.5	58	5 095092	095u92 eriocheir s
29	49	12.5	59	5 09TW8	09twf8 anemonia su
30	48.5	12.4	29	10 09S8D2	09s8d2 cucumis mel
31	48.5	12.4	31	6 077625	077625 bos taurus
32	48.5	12.4	41	4 090DP7	090dp7 homo sapien
33	48.5	12.4	55	17 09HJ78	09hj78 thermoplasm
34	47.5	12.1	37	11 09QX87	09qx87 rattus norv
35	47.5	12.1	57	5 09N9H2	09n9h2 venerupis (
36	47.5	12.1	61	4 09BXG3	09bxg3 homo sapien
37	47.5	12.1	61	6 018780	018780 ovis aries
38	47	12.0	41	13 09PVG6	09pv66 xiphophorus
39	47	12.0	47	5 09N9H0	09n9h0 venerupis p
40	47	12.0	47	5 09VIY8	09viy8 drosophila
41	47	12.0	56	2 093GH5	093gh5 bacillus su
42	47	12.0	61	5 09BJK0	09bjk0 bacterocera
43	46.5	11.9	67	5 095P50	095p50 helix pomat
44	46.5	11.9	67	13 090793	090793 gallus gall
45	46	11.8	48	11 060413	060413 cricetus cr

ALIGNMENTS

RESULT 1

077419 ID 077419 PRELIMINARY; PRT; 62 AA.
AC 077419;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Chymotrypsin/elastase inhibitor-1 (Fragment).
GN ASC/E-1.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea.
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98297373; PubMed=9635450;
RA Lu C.-, Nguyen T., Morris S., Hill D., Sakanari J.A.;
RT "Anisakis simplex: mutational bursts in the reactive site centers of
RT serine protease inhibitors from an ascarid nematode.";
RL Exp. Parasitol. 89:257-261(1998).
DR EMBL; P07851; LEAI.
DR HSSP; P07851; LEAI.
DR InterPro: IPR002919; TIL_Cysrich.
DR Pfam: PF01826; TIL; 1.
FT NON_TER 1
FT NON_TER 62 62
SQ SEQUENCE 62 AA; 6794 MW; 154CE25A375E0B2C CRC64;
Query Match 22.6%; Score 88.5; DB 5; Length 62;
Best Local Similarity 33.3%; Pred. No. 0.00022;
Matches 21; Conservative 9; Mismatches 26; Indels 7; Gaps 4;

QY 8 GRGKCPSEIFSRCDGRCQFCC--PNVVPKPLCIKICAPGCVCR--LGYLRNKKKVCVPR 63
! : !
Db 1 GQERCGPNEVWTECTG-CEMKCGDPENTPCPLMCR--RPSCECSGRCGMRRTNDGKCLPA 57
QY 64 SKC 66
! ! !
Db 58 SQC 60

[illegible]

Db 10 JLCVSVSV-----CVSVSVSVCLVYMR-----AHMPMSK 42

RESULT 6

Q8TDN3 ID Q8TDN3 PRELIMINARY; PRT; 61 AA.

AC Q8TDN3; 21, Created

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DE Metallothionein 1M.

GN MTM.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ON NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang J., Yu L., Zhao S.;

RT "Cloning of a novel member of the MT gene family-MT1M.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF348571; AAL83902.1;

FT SEQUENCE 61 AA; 6110 MW; 754753ED088411CE CRC64;

SQ

Query Match 15.3%; Score 60; DB 4; Length 61;

Best Local Similarity 39.1%; Pred. No. 1.4;

Matches 18; Conservative 5; Mismatches 17; Indels 6; Gaps 5;

QY 10 GKPSNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVCLYLRN 55

Db 17 GSCTCKE--CKCYS-CKKSCCCCPVG-CAK-CAHGCVCVK-GTLEN 56

RESULT 7

Q9BSK9 ID Q9BSK9 PRELIMINARY; PRT; 63 AA.

AC Q9BSK9; 17, Created

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DE Hypothetical 5.9 kDa protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ON NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC004969; AAH04969.1;

KW Hypothetical protein.

SQ SEQUENCE 63 AA; 5927 MW; 8DF7AB2529D8628C CRC64;

Query Match 15.2%; Score 59.5; DB 4; Length 63;

Best Local Similarity 29.4%; Pred. No. 1.7;

Matches 15; Conservative 4; Mismatches 19; Indels 13; Gaps 2;

QY 2 GFEGGLGRCGKCPSEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVCLYLRN 46

Db 15 GFSGVLGGRCQCG-----GSCRSGLGPPASSLLCCKPASGPMVCLGCG 58

RESULT 8

Q98TCO ID Q98TCO PRELIMINARY; PRT; 48 AA.

AC Q98TCO; 17, Created

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DE Metallothionein (MT) (Fragment).

GN MT.

OS Seriola quinqueradiata (Five-ray yellowtail).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;

OC Carangidae; Seriola.

ON NCBI_TaxID=8161;

RN [1]

RP SEQUENCE FROM N.A.

RA Putami K., Maita M.;

RT "Molecular cloning of metallothionein cDNA from yellow tail (Seriola

quinqueradiata) for monitoring of environmental pollution.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE

RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.

DR EMBL; AB059742; BAB41015.1;

DR HSSP; P02795; 1MHU.

DR InterPro; IPR003019; Metallothion.

DR InterPro; IPR000006; Metallothion_vert.

DR Pfam; PF00131; metalthio; 1.

DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.

KW Chelation; Metal-binding; Metal-thiolate cluster.

FT NON_TER 1

FT NON_TER 48

SQ SEQUENCE 48 AA; 4732 MW; 17ADD345E376867 CRC64;

Query Match 15.1%; Score 59; DB 13; Length 48;

Best Local Similarity 35.2%; Pred. No. 1.5;

Matches 19; Conservative 3; Mismatches 18; Indels 14; Gaps 4;

QY 7 GGRGKCPSEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVCLYLRNKKVC 60

Db 8 GGSCTC-INCSTCTCKKSCCPCCPSG-----CTK-CASGCVG-----KGKTC 47

RESULT 9

Q9BQN2 ID Q9BQN2 PRELIMINARY; PRT; 58 AA.

AC Q9BQN2; 17, Created

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Metallothionein (MT).

GN DJ61404.6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ON NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Clark G.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE

RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.

DR EMBL; AL121753; CAG34280.1;

DR HSSP; P02795; 1MHU.

DR InterPro; IPR001396; Echnd_Metthion4.

DR InterPro; IPR003019; Metallothion.

DR InterPro; IPR000006; Metallothion_vert.

DR Pfam; PF001008; Metalthion_misc.

DR Pfam; PF00131; metalthio; 1.

DR PRINTS; PR00873; MTMOLLUSC.

DR PRINTS; PR00875; MTMOLLUSC.

DR PROSITE; PS00860; MTVERTBRATE.

DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.

KW Chelation; Metal-binding; Metal-thiolate cluster.

SQ SEQUENCE 58 AA; 5760 MW; D386349E68A2DA3B CRC64;

Query Match 15.0%; Score 58.5; DB 4; Length 58;

Best Local Similarity 37.5%; Pred. No. 2.1;

Matches 15; Conservative 5; Mismatches 15; Indels 5; Gaps 4;

QY 10 GKPSNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVCLYLRN 49

RESULT 14


```

Q8TDC4
ID Q8TDC4 PRELIMINARY; PRT; 61 AA.
AC Q8TDC4;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Metallothionein 1Y.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Weyenbergh J., De Ley M.;
RT "Identification of molecular targets for zinc regulation of human
RL monocytic metabolism."
DL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF479613; AAL85413.1; -
SQ SEQUENCE 61 AA; 6179 MW; 705F96E8B99D1CE CRC64;

Query Match 13.9%; Score 54.5; DB 4; Length 61;
Best Local Similarity 38.5%; Pred. No. 7.5;
Matches 15; Conservative 4; Mismatches 15; Indels 5; Gaps 4;

QY 10 GKCPSEIFSRCDRCORFCPNVVPKPLCIKICAPGCVC 48
Db 17 GSKCKE--CKCTS--CKKSCCPCPGV--CAK-CAHGCVC 50

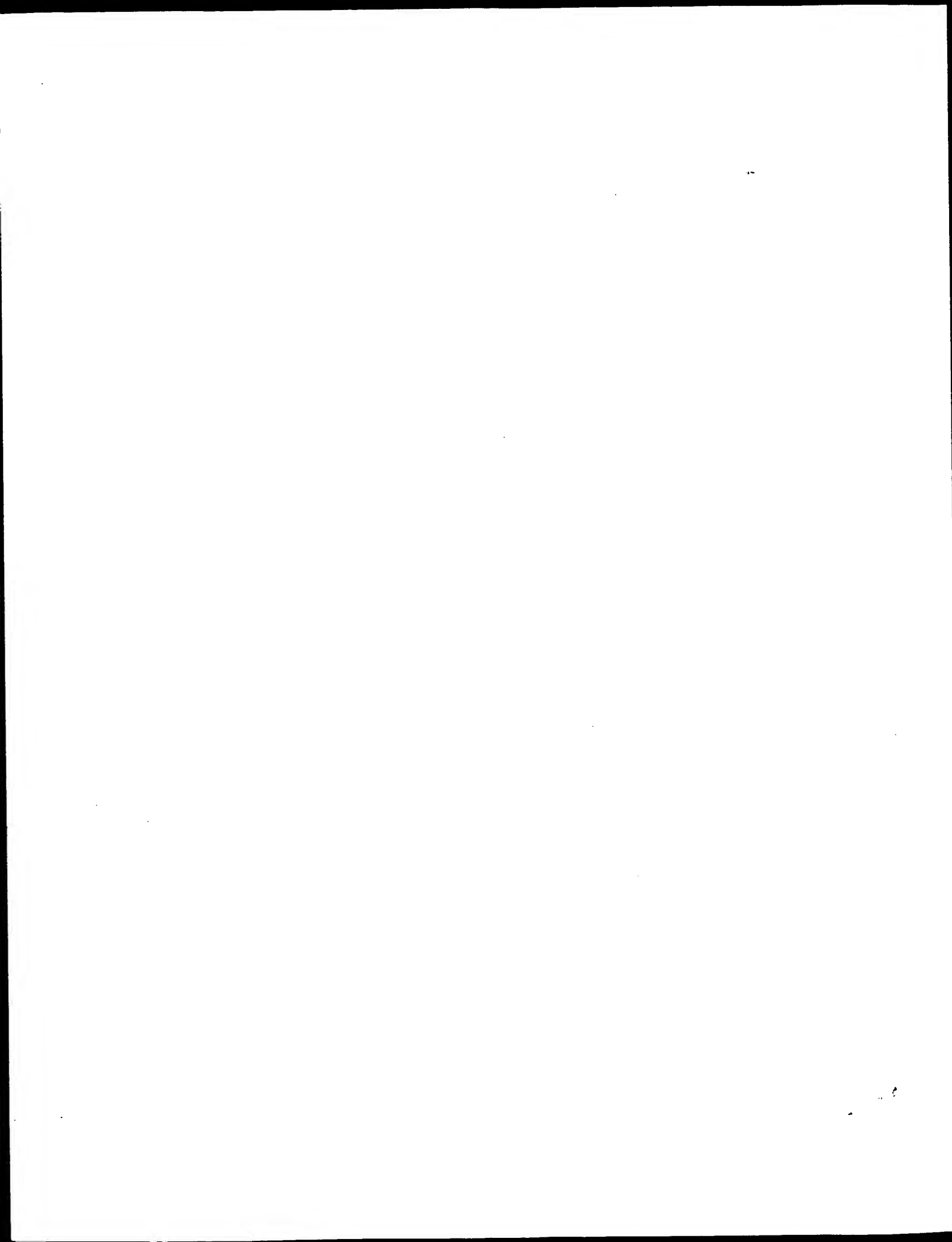
RESULT 15
Q9PVG7
ID Q9PVG7 PRELIMINARY; PRT; 39 AA.
AC Q9PVG7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase Xmrk (Fragment).
GN XMRK.
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8083;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SR;
RX MEDLINE=99126443; PubMed=9927468;
RA Gutbrod H., Scharl M.;
RT "Intragenic sex-chromosomal crossovers of Xmrk oncogene alleles affect
RT pigment pattern formation and the severity of melanoma in
RT Xiphophorus."
RL Genetics 151:773-783(1999).
DR EMBL: AF092693; AAD03714.1; -
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
KW Kinase.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4195 MW; 42303BA05F6DB557 CRC64;

Query Match 13.6%; Score 53; DB 13; Length 39;
Best Local Similarity 42.9%; Pred. No. 7.7;
Matches 12; Conservative 3; Mismatches 9; Indels 4; Gaps 2;

QY 21 CDGRCORFCPNVVPKPL--CIKICAPGC 46
Db 5 CAEQCNRRCRG--PKPIDCCNEHCAGGC 30

```

Search completed: February 26, 2003, 15:12:38
Job time : 92 secs



GenCore version 5.1.3
Copyright. (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:05:47 ; Search time 35 Seconds
(without alignments)
255.080 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGGGGLGGKCPSEIFSR.....CRLGLRNKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 497633

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	391	100.0	67	ABB08330	Bee venom protein
2	82	21.0	61	AAV30434	Mature nematode ex
3	82	21.0	61	AA15319	N. american nemat
4	68	17.4	58	AAV30433	Mature nematode ex
5	63.5	16.2	59	AAV57812	Trout metallothion
6	63.5	16.2	59	ABP32832	Human ORF1805 prot
7	62.5	16.0	62	ABP77804	Consensus sequence
8	62.5	16.0	63	AA021321	Arabidopsis thalia
9	62	15.9	43	AA060286	Human factor IX (h
10	62	15.9	61	AAV57822	Rabbit liver metal

11	61.5	15.7	67	22	AA061455	Metallothionein do
12	61	15.6	47	20	AAV04648	Factor IXa catalyt
13	61	15.6	56	21	AA044778	Human secreted pro
14	61	15.6	66	22	AAU45263	Propionibacterium
15	60	15.3	60	23	ABP01364	Human ORFX protein
16	59.5	15.2	61	22	AAU28133	Novel human secret
17	59	15.1	61	19	AAW61601	Human metallothion
18	58.5	15.0	39	20	AAV02082	KTX peptide used t
19	58.5	15.0	61	22	AAW40263	Human polypeptide
20	58	14.8	56	22	AAU48798	Propionibacterium
21	57.5	14.7	67	22	AAW79414	Human protein SEQ
22	57	14.6	10	21	AAV69218	N-terminal sequenc
23	57	14.6	40	5	AAV40692	Sequence encoded b
24	57	14.6	40	5	AAV40220	Partial sequence o
25	56.5	14.5	55	22	AAW39148	Human polypeptide
26	56.5	14.5	60	14	AAV40209	Sequence of human
27	56.5	14.5	60	21	AAV82332	Human metallothion
28	56.5	14.5	61	21	AAV12587	Human metallothion
29	56.5	14.5	61	21	AAV82331	Human metallothion
30	56.5	14.5	61	22	AAW78430	Human protein SEQ
31	56.5	14.5	61	23	ABP09810	Amino acid sequenc
32	56	14.3	66	23	ABP02424	Human ORFX protein
33	55.5	14.2	46	21	AAV12508	Beta-cellulin mute
34	55.5	14.2	47	21	AAV12507	Beta-cellulin mute
35	55.5	14.2	47	21	AAV12516	Beta-cellulin mute
36	55.5	14.2	48	21	AAV12515	Beta-cellulin mute
37	55.5	14.2	48	21	AAV12517	Beta-cellulin mute
38	55.5	14.2	48	21	AAV12518	Beta-cellulin mute
39	55.5	14.2	53	21	AAV12521	Beta-cellulin mute
40	55.5	14.2	60	21	AAW49082	Metallothionein fr
41	55.5	14.2	63	21	AAV57811	Chicken metallothi
42	55.5	14.2	67	23	ABP09671	Human ORFX protein
43	55	14.1	51	6	AAV50547	Protein analogue e
44	55	14.1	66	22	AAU55134	Propionibacterium
45	54.5	13.9	48	21	AAV12514	Beta-cellulin mute

ALIGNMENTS

RESULT 1

ABB08330
ID ABB08330 standard; protein; 67 AA.

AC ABB08330;

DT 18-JUN-2002 (first entry)

XX Bee venom protein Api m 6.01.

DE Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;
KW Bee venom hypersensitivity; antibody; protein purification; Api m 6.01;
KW immunotherapy; allergen.

XX Apis sp.

XX WO200188085-A2.

XX 22-NOV-2001.

PF 16-FEB-2001; 2001WO-IB01736.

XX 18-FEB-2000; 2000US-0506978.

XX (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

PI Spertini F;

XX WPI; 2002-082988/11.

PT New bee venom polypeptides, useful for modulating immune responses e.g.
PT in individual hypersensitive to the venom and for identifying
PT individual at risk for bee venom hypersensitivity

XX Example 2; Page 26; 32pp; English.

CC The present sequence is that of one the four isoforms of Api m 6.

CC designated Api m 6.01. The sequence represents the central amino acid

CC sequence shared by all four isoforms (see ABB08331, ABB08332 and

CC ABB08333). The specification describes a substantially pure polypeptide,

CC Api m 6, derived from bee venom and found in four isoforms. The proteins

CC of the invention have immunosuppressant activity and may form the basis

CC of a vaccine. Api m 6 is useful for modulating an immune response, i.e.

CC as an allergen for immunotherapy. The protein is useful for identifying

CC an individual at risk for bee venom hypersensitivity. The method

CC comprises administering Api m 6 to the individual and measuring an immune

CC response raised, where a detectable immune response indicates that the

CC individual is at risk for bee venom hypersensitivity. Antibodies specific

CC for Api m 6 are useful for purifying the protein.

XX Sequence 67 AA;

Query Match 100.0%; Score 391; DB 23; Length 67;

Best Local Similarity 100.0%; Pred. No. 8e-32;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPNEIFSRCDGRCPNVVPKPLCIKICAPGCVCRGLYLNRKKVC 60

Db 1 GFGGLGGRGKCPNEIFSRCDGRCPNVVPKPLCIKICAPGCVCRGLYLNRKKVC 60

Qy 61 VPRSKCG 67

Db 61 VPRSKCG 67

RESULT 2

AAV30434

ID AAV30434 standard; Protein; 61 AA.

XX AAV30434;

XX AAV30434;

DT 15-NOV-1999 (first entry)

DE Mature nematode extracted anticoagulant protein NamNAP.

XX Nematode extracted anticoagulant protein; NAP; anticoagulant;

KW serine protease inhibitor; NAP domain; factor VIIa/TF.

XX Necator americanus.

OS US5955294-A.

PN 21-SEP-1999.

XX 19-APR-1996; 96US-0634641.

XX 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

XX Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;

PI Vlasuk GP;

XX WPI; 1999-539569/45.

XX Screening an isolated protein for Nematode-extracted Anticoagulant

PT protein domains

XX Disclosure; Columns 143-144; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant

CC protein (NAP). The protein has activity as an anticoagulant and/or serine

CC protease inhibitor. The protein contains at least one NAP domain which

CC has selective inhibitory activity for factor VIIa/TF. The specification

CC describes a method for screening an isolated protein at least one domain

CC for factor VIIa/TF selective inhibitory activity. The method comprises

CC determining the time to clotting effected by a concentration of the

CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo

CC activated partial thromboplastin time (aPTT) assay; calculating

CC prolongation of clotting effected by the isolated protein in each of

CC the PT and aPTT assay, with respect to a baseline clotting value for

CC each assay, where prolongation of clotting is calculated as fold

CC elevation of clotting time relative to a baseline clotting value, where

CC a doubling of clotting time is deemed a two-fold elevation; and

CC calculating a PT to aPTT prolongation ratio, where a ratio at least

CC one is indicative of factor VIIa/TF inhibitory activity. The method is

CC useful for determining if a protein has factor VIIa/TF inhibitory

CC activity.

XX Sequence 61 AA;

Query Match 21.0%; Score 82; DB 20; Length 61;

Best Local Similarity 30.9%; Pred. No. 0.37;

Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;

Qy 12 CFSNEIFSRCDGRCPNVVPKPLCIKICAPGCV-----CRLGLYLRNKKVCV 61

Db 4 CPANEWRECGTPCEPKCNQMP-----DICTMNCIVDCQCKEGYKRHETKGCL 53

RESULT 3

AAV15319

ID AAV15319 standard; Protein; 61 AA.

XX AAV15319;

XX AAV15319;

DT 19-DEC-2000 (first entry)

DE N. americana nematode-extracted anticoagulant protein NamNAP.

XX Nematode-extracted anticoagulant protein; NamNAP; blood clotting;

KW canine hookworm; thrombosis; vaccine.

XX Necator americanus.

OS US6087487-A.

PN 11-JUL-2000.

XX 12-FEB-1999; 99US-0249451.

XX 17-OCT-1995; 95WO-US13231.

PR 17-APR-1997; 97US-0809455.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

XX (CORV-) CORVAS INT INC.

PA Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;

XX WPI; 2000-531359/48.

DR N-PSDB; AAA73383.

XX New cDNA molecule encoding a protein having factor Xa inhibitory

PT activity for preventing and treating blood clotting disorders,

PT comprises nematode-extracted anticoagulant protein domains

XX Disclosure; Fig 16; 197pp; English.

PS Disclosure; Column 3; 9pp; English.

XX The present sequence represents mammalian metallothionein. It is
 CC used to produce chimeric proteins of the invention. The specification
 CC describes a recombinant bifunctional streptavidin-metallothionein
 CC chimeric protein. This protein is produced by introducing into a host
 CC cell nucleic acid encoding a bifunctional fusion protein having a
 CC streptavidin and a metallothionein moiety, and incubating the cell to
 CC express the fusion protein. The streptavidin moiety consists of residues
 CC 16-133 of mature streptavidin. The chimeric protein is used to
 CC incorporate heavy metal ions into biological materials containing biotin,
 CC or to remove heavy metal ions from the biological material. Specific uses
 CC include loading cancerous tissue with heavy metal ions for imaging of
 CC tumour cells and radiotherapy, and labeling DNA and proteins for
 CC detection on gels or blots by surface scanning mass spectrometry.

XX Sequence 62 AA;

Query Match 16.0%; Score 62.5; DB 23; Length 62;
 Best Local Similarity 37.3%; Pred. No. 32;
 Matches 19; Conservative 2; Mismatches 23; Indels 7; Gaps 3;

OY 1 GCGGLGGRGKPSNEIFSRCDGRQPCPNVVPKLCIKICAPGCVCLRG 51
 DB 10 GGSCTAGSCCKCKCKCTS-CKKSCCSCP-----VGCAK-CAQCGCVCKAG 53

RESULT 8

AA021321
 ID AA021321 standard; Protein; 63 AA.

XX AA021321;

XX 05-AUG-2002 (first entry)

XX Arabidopsis thaliana KCP-like protein, SEQ ID NO 76.

XX Antimicrobial; transgenic; plant; potato snakin antimicrobial protein;
 KW GAS44; GAS45; GST1 homologue; lysine- and cysteine- rich peptide;
 KW KCP-like polypeptide; modulating; disease resistance.

OS Arabidopsis thaliana.

XX WO200222821-A2.

XX 21-MAR-2002.

XX 13-SEP-2001; 2001WO-US28429.

XX 13-SEP-2000; 2000US-232569P.

XX 11-SEP-2001; 2001US-0950933.

XX (PTON-) PIONEER HI-BRED INT INC.

XX Simmons CR, Navarro Acevedo PA;

XX WPI; 2002-425842/45.

XX New polynucleotide encoding lysine- and cysteine-rich peptides-like
 PT polypeptide useful for modulating the polypeptide level in a plant
 PT cell, enhancing disease resistance -

XX Disclosure; Page 154; 163pp; English.

XX The invention relates to an isolated polynucleotide encoding a
 CC polypeptide which is related to potato snakin antimicrobial protein and
 CC GAS44 or GAS45 or GST1 homologues, which is referred to as lysine- and
 CC cysteine- rich peptides (KCP)-like polypeptide, having a nucleotide
 CC sequence from 36 residues of defined base pairs, given in the
 CC specification. A recombinant expression cassette comprising the isolated
 CC polynucleotide of the invention is useful for modulating the level of
 CC (KCP)-like polypeptides in a plant cell, where the level of (KCP)-like
 CC polypeptides is increased, and disease resistance is enhanced. This

CC sequence represents a protein relating to the KCP-like proteins of the
 CC invention.

XX Sequence 63 AA;

Query Match 16.0%; Score 62.5; DB 23; Length 63;
 Best Local Similarity 37.8%; Pred. No. 33;
 Matches 17; Conservative 4; Mismatches 19; Indels 5; Gaps 3;

OY 20 RCDGRQRCPCPNV-VPKP---LCIKICAPGCVCLRGVLRNKKKVC 60
 DB 4 QCGGQCTRRCSNTYHKPCMFECQKCAK-CLCVPPPTGYGNKQVC 47

RESULT 9

AAB60286
 ID AAB60286 standard; Protein; 43 AA.

XX AAB60286;

XX 30-MAR-2001 (first entry)

XX Human factor IX (hFIX) exon-encoded fragment, SEQ ID NO:9.

XX Age-related gene regulation; liver-specific; gene expression;
 KW human factor IX; hFIX; AE5'; AE3'; age-regulatable expression construct;
 KW antisense therapy; gene therapy; thrombosis; cardiovascular disease;
 KW diabetes; Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
 KW osteoarthritis; dementia.

XX Homo sapiens.

XX WO200075279-A2.

XX 14-DEC-2000.

XX 06-JUN-2000; 2000WO-US15728.

XX 09-JUN-1999; 99US-0328925.

XX (UNMI) UNIV MICHIGAN.

XX Kurachi K, Kurachi S;

XX WPI; 2001-061708/07.

XX N-PSDB; AAF54018.

XX New regulatory elements that control age-related gene expression,
 PT useful in gene therapy and for reducing Factor IX expression -

XX Disclosure; Fig 8C; 225pp; English.

XX The invention relates to nucleic acid sequences which regulate gene
 CC expression in an age-related manner and/or in a liver-specific manner.
 CC The invention identifies regions of the human factor IX (hFIX) gene, and
 CC a region of the human protein C (hPC) gene, which are age-related
 CC regulatory sequences. The hFIX age-related regulatory sequences are
 CC designated AE5' (AAF54016) and AE3' (AAF54017) and are found in the 5'
 CC UTR (at position 2164-2165 of AAF54018) and 3' UTR (at position
 CC 34383-35655 of AAF54018) respectively. These elements act synergistically
 CC to increase hFIX levels over the lifespan of an individual; however, they
 CC can independently exert effects on hFIX mRNA in an age-related manner,
 CC with AE5' acting to stabilise hFIX mRNA, and AE3' acting to increase hFIX
 CC mRNA levels, over time. AE5' also directs liver-specific expression. The
 CC hPC gene age-related regulatory sequence is found in the 5' UTR
 CC (AAF54081), and contains two PEA-3 (polyoma virus activator 3) elements
 CC 5'-GAGGAAGA-3' and 5'-CAGGAAG-3'. The age-related regulatory sequences of
 CC the invention, along with their homologues, variants and fragments, may
 CC be used in the construction of recombinant expression vectors for the
 CC expression of a desired sequence in an age-related fashion in a host
 CC cell. Preferred target genes for expression in such age-regulatable
 CC expression vectors include those encoding proteins involved in blood
 CC coagulation (e.g., the pro-coagulants factor IX and factor VIII, and the

CC anti-coagulants protein C and antithrombin III), human
 CC alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
 CC luciferase. Preferred promoters for use in such age-regulatable
 CC expression vectors include the human factor IX promoter, the T7 promoter,
 CC the T3 promoter and the SP6 promoter. The expression vectors of the
 CC invention may be used in gene therapy to provide age-related and/or
 CC liver-specific expression of target genes. Age-regulatable constructs may
 CC be used in the treatment of such age-related conditions such as
 CC thrombosis, cardiovascular disease, diabetes, Alzheimer's disease,
 CC Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
 CC Specifically, they may be used to express factor IX antisense mRNA in the
 CC treatment of thrombotic conditions associated with the natural
 CC age-related rise in factor IX expression. Transgenic cells or animals
 CC that contain vectors of the invention are useful as models of these
 CC diseases, in screening for potential therapeutic agents and for studying
 CC normal processes such as aging and gene expression. Fragments and
 CC homologues of age-related regulatory sequences, are useful as probes to
 CC detect, isolate or identify other such sequences in samples. The present
 CC sequence represents a fragment of hFIX.

XX Sequence 43 AA;

Query Match 15.9%; Score 62; DB 22; Length 43;
 Best Local Similarity 33.3%; Pred. No. 26;
 Matches 14; Conservative 5; Mismatches 15; Indels 8; Gaps 2;

Qy 22 DGRGRCFNPVVPKPLCICAPGCVRLGY-LRNKKKVCVP 62
 Db 8 NGRGRCFNPVVPKPLCICAPGCVRLGY-LRNKKKVCVP 42

RESULT 10
 AAY57822

ID AAY57822 standard; protein; 61 AA.

XX AAY57822;

XX 22-MAR-2000 (first entry)

XX Rabbit liver metallothionein Class II amino acid sequence.

XX Metallothionein; metal recovery; remediation; heavy metal;
 KW precious metal; phytochelatin; green algae; Chlamydomonas reinhardtii.

XX Oryctolagus cuniculus.

XX WO9960838-A1.

XX 02-DEC-1999.

XX 28-MAY-1999; 99WO-US12007.

XX 28-MAY-1999; 98US-0087374.

XX (OHIS) UNIV OHIO STATE RES FOUND.

XX Sayre RT, Traina SJ;

XX WPI; 2000-086646/07.

XX Novel method for metal recovery, remediation and separation -

XX Disclosure; Fig 1; 86pp; English.

XX The present invention describes a transgenic algal cell (I) of the
 CC genus Chlamydomonas comprising reproductive genetic material comprising
 CC a nucleotide sequence capable of expressing chicken type I
 CC Metallothionein. Also described is a method of removing metal from
 CC an aqueous medium containing at least one dissolved or suspended
 CC metal. The transgenic algae are used for the selective separation of
 CC metals, particularly the separation of precious and desirable metals
 CC such as gold and uranium, from other metals such as cadmium, zinc and
 CC copper. The method can be used to facilitate the selective recovery of

CC precious and rare metals from mineral sources where aqueous media can
 CC be used, such as in natural surface water flows, ground water and where
 CC water may be introduced. The method is suitable for well-drilling,
 CC soil and water remediation arts, mining fields, and industrial
 CC engineering. The present sequence represents a Class II metallothionein
 CC given in the present invention.

XX Sequence 61 AA;

Query Match 15.9%; Score 62; DB 21; Length 61;
 Best Local Similarity 38.5%; Pred. No. 36;
 Matches 15; Conservative 2; Mismatches 10; Indels 12; Gaps 3;

Qy 11 KCPSEIFSRCDGRGRCFNPVVPKPLCICAPGCVCR 49
 Db 25 KCTPS-----CKKSCCSCP-----PGCAK-CAAGCICK 51

RESULT 11

AAB61455

ID AAB61455 standard; protein; 67 AA.

XX AAB61455;

XX 04-APR-2001 (first entry)

XX Metallothionein domain consensus.

XX TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
 KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
 KW pancreatic; skeletal; muscle.

XX Synthetic.

XX WO200100672-A1.

XX 04-JAN-2001.

XX 29-JUN-2000; 2000WO-US18184.

XX 29-JUN-1999; 99US-0342687.

XX (MILL-) 'MILLENNIUM PHARM INC.

XX Holtzman DA, Barnes TM, Fraser CC, Sharp JD;

XX WPI; 2001-050127/06.

XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
 CC MANGO 245 proteins, useful in the treatment of inflammatory diseases
 CC (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
 CC disorders (e.g. jaundice) -

XX Disclosure; Fig 13; 262pp; English.

XX The present invention relates to cDNAs encoding TANGO 244,

CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.

CC The nucleic acids, proteins and protein modulators are useful for

CC treating colonic disorders, inflammatory diseases, tumors,

CC renal disorders, liver disorders, lung disorders, autoimmune diseases,

CC allergic diseases, cardiovascular diseases, brain disorders,

CC degenerative diseases placental, pancreatic, skeletal and muscle

XX disorders.

XX Sequence 67 AA;

Query Match 15.7%; Score 61.5; DB 22; Length 67;

Best Local Similarity 31.7%; Pred. No. 43;

Matches 19; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

Qy 1 GFGGLGGRGKCPSEIFSRCDGRGRCFNPVVPKPLCICAPGCVRLGYNKKVC 60

Db 11 GGCCTGCTGCKCKCKCTSCCKSCCCCPAGCSK-----CAGGCVCCKGGGAASETSKC 63


```

RESULT 12
ID AAY04648
XX AC AAY04648 standard; peptide; 47 AA.
XX AC AAY04648;
XX DT 22-JUN-1999 (first entry)
XX DE
XX DE Factor IXa catalytic and interacting domains.
XX KW Receptor; catalytic domain; Factor IXa; Factor Xa; tissue factor; angina;
XX KW blood clotting disorder; thrombosis; restenosis; myocardial infarction;
XX KW reoclosure; cerebrovascular disease; hypercoagulability; anticoagulant;
XX KW peripheral arterial occlusive disease; pulmonary embolism; cyclic.
XX OS Homo sapiens.
XX FH
XX FH Key Location/Qualifiers
XX FT Disulfide-bond 8..19
XX FT Domain 10..12
XX FT /note= "Factor VIIa interacting domain"
XX FT Disulfide-bond 13..29
XX FT Domain 14..16
XX FT /note= "catalytic domain"
XX FT Disulfide-bond 31..44
XX FT
XX FT W09913062-A1.
XX PN
XX PD 18-MAR-1999.
XX PF 08-SEP-1998; 98WO-GB02700.
XX PR 09-SEP-1997; 97GB-0019157.
XX PR (MATT/) MATTHEWS D P.
XX PA (NYCO-) NYCOMED IMAGING AS.
XX PI Fischer PM, Sakariassen KS;
XX PI WPI; 1999-215061/18.
XX DR
XX DR New anticoagulant compounds
XX PT
XX PS Disclosure; Fig 1; 50pp; English.
XX CC
XX CC Peptides AAY04625-Y04647 represent claimed compounds which are capable
XX CC of interacting with an internal receptor in the catalytic domain of
XX CC Factor IXa (FIXa) or Factor X (FX) defined by the residues Ile290,
XX CC Ala291, Asp292, Tyr293, Thr294, Glu374 and Phe378 of FIXa, and Leu300,
XX CC Pro301, Glu302, Trp305, Ala306, Lys385 and Phe389 of FXa, or the ligand
XX CC defined by residues Cys95-Cys99 of FIXa or Cys96-Cys100 of FXa. The
XX CC compounds can be used to prevent the formation of a functional
XX CC FVIIa/FIXa or FVIIa/FXa complex, so can be used to combat or prevent
XX CC blood clotting disorders, e.g. thrombosis (particularly vascular
XX CC thrombosis or deep vein thrombosis), acute myocardial infarction,
XX CC restenosis, angina, reoclosure, cerebrovascular disease, peripheral
XX CC arterial occlusive disease, hypercoagulability or pulmonary embolism.
XX CC They can also be used to prevent the occurrence of blood clotting
XX CC disorders caused by e.g. grafting surgery, vessel wall potency
XX CC restoration or sepsis.
XX CC
XX CC Sequence 47 AA;
XX CC
Query Match 15.6%; Score 61; DB 20; Length 47;
Best Local Similarity 34.1%; Pred. No. 36;
Matches 14; Conservative 4; Mismatches 15; Indels 8; Gaps 2;

QY 23 GRCQRCFNPVVKPLCIKICAPGCVCRIGY-LRNKKKVCVP 62
||||:| | | | | | | | | | | | | | | | | |
Db 13 GRCEQFCKNSADNKVV-----CSCTEGYRLAENQKSCEP 46

```

RESULT 13

```

AAB44778
XX ID AAB44778 standard; Protein; 56 AA.
XX AC AAB44778;
XX DT 12-FEB-2001 (first entry)
XX DE
XX DE Human secreted protein sequence encoded by gene 17 SEQ ID NO:77.
XX KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
XX KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
XX KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
XX KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
XX KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
XX KW cerebrovascular disorder; nervous system disorder; ocular disorder;
XX KW wound healing; skin aging; food additive; preservative.
XX OS Homo sapiens.
XX PN W0200058336-A1.
XX PD 05-OCT-2000.
XX PF 23-MAR-2000; 2000WO-US07726.
XX PR 26-MAR-1999; 99US-0126597.
XX PR 07-JAN-2000; 2000US-0174877.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM, Komatsoulis G;
XX PI WPI: 2000-6023355/57.
XX DR N-PSDB; AAC79815.
XX PT
XX PT Nucleic acid encoding human secreted proteins, used to treat, prevent,
XX PT ameliorate or diagnose medical conditions such as cancer, and
XX PT autoimmune diseases.
XX PS Claim 11; Page 362; 391pp; English.
XX CC
XX CC The polynucleotide sequences given in AAC79799 to AAC79848 encode the
XX CC human secreted proteins given in AAB44762 to AAB44811. AAB44812 to
XX CC AAB44829 represent human secreted polypeptide sequences and proteins
XX CC homologous to them, which are used in the exemplification of the present
XX CC invention. Human secreted proteins have activities based on the tissues
XX CC and cells the genes are expressed in. Examples of activities are:
XX CC immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
XX CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;
XX CC neuroprotective; antibacterial; virucide; fungicide; and
XX CC ophthalmological. The polynucleotides and polypeptides can be used to
XX CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
XX CC in diagnosing a pathological condition or susceptibility to a
XX CC pathological condition. Disorders which are diagnosed or treated include
XX CC autoimmune diseases, hyperproliferative disorders, cardiovascular
XX CC disorders, cerebrovascular disorders, angiogenesis, nervous system
XX CC disorders, infections caused by bacteria, viruses and fungi and ocular
XX CC disorders. The polypeptides can also be used to aid wound healing and
XX CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
XX CC maintain organs before transplantation, for supporting cell culture of
XX CC primary tissues, to regenerate tissues and in chemotaxis. The
XX CC polypeptides can also be used as a food additive or preservative to
XX CC increase or decrease storage capabilities. AAC79790 to AAC79798 and
XX CC AAB44761 represent sequences used in the exemplification of the present
XX CC invention.
XX CC
XX CC Sequence 56 AA;
XX CC
Query Match 15.6%; Score 61; DB 21; Length 56;
Best Local Similarity 29.8%; Pred. No. 42;

```

Matches 14; Conservative 3; Mismatches 18; Indels 12; Gaps 2;		Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;	
QY	21 CDGRCQRCPCNVVVKPLIKICAPCCVCLGLYLRNKKKVCVPSKCG 67	QY	2 GFGLGGRGKC----PSNEIFSRCDGRCPNVPKPLIKICAPGCVCLGLYLRNKKK 58
DB	4 CVQACVHMC-----LCVHVCASVHVCEAYVH-----CVREACG 38	DB	7 GRGLISAGKGRPSDQ--RRCSAOCGGLPGSRPVSRCVR-----CIC--CSLRASPL 57
RESULT 14		RESULT 15	
AAU45263		ABP01364	
ID	AAU45263 standard; Protein: 66 AA.	ID	ABP01364 standard; Protein: 60 AA.
AC	AAU45263;	XX	ABP01364;
DT	27-FEB-2002 (first entry)	XX	25-JUN-2002 (first entry)
DE	Propionibacterium acnes immunogenic protein #6159.	XX	Human ORFX protein sequence SEQ ID NO:2710.
XX	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;	XX	Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;	KW	hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;	KW	degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW	dermatological; osteopathic; neuroprotectant.	KW	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX	Propionibacterium acnes.	KW	hypertension; hypothyroidism; cholesterol ester storage disease;
OS	WO200181581-A2.	KW	immune deficiency; immune disorder; infectious disease;
PN	01-NOV-2001.	KW	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
PD		KW	myasthenia gravis.
PF	20-APR-2001; 2001WO-US12865.	XX	Homo sapiens.
XX		OS	WO200192523-A2.
XX	21-APR-2000; 2000US-199047P.	XX	06-DEC-2001.
PR	02-JUN-2000; 2000US-208841P.	XX	29-MAY-2001; 2001WO-US10836.
PR	07-JUL-2000; 2000US-216747P.	XX	30-MAY-2000; 2000US-206132P.
XX		PR	29-AUG-2000; 2000US-228716P.
XX	(CORI-) CORIXA CORP.	XX	(CURA-) CURAGEN CORP.
PA		XX	Shimkets RA, Leach MB;
XX	Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;	PI	WPI; 2002-106308/14.
PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;	DR	N-PSDB; ABN17116.
PI		DR	
XX	WPI; 2001-616774/71.	XX	Novel human polypeptides and polynucleotides useful for diagnosing,
DR	Propionibacterium acnes polypeptides and nucleic acids useful for	XX	preventing and treating cardiovascular disease, neurodegenerative,
DR	vaccinating against and diagnosing infections, especially useful for	XX	hyperproliferative disorders and autoimmune disorders
DR	treating acne vulgaris -	XX	Disclosure; SEQ ID 2710; 1037pp; English.
XX	Example 1; SEQ ID NO 6458; 1069pp; English.	XX	The present invention describes substantially purified human proteins
PS	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic	CC	(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX	polypeptides. The proteins and their associated DNA sequences are used in	CC	in the specification). ABN15762 to ABN27252 encode the human ORFX
CC	the treatment, prevention and diagnosis of medical conditions caused by	CC	proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,	CC	treating or preventing a pathology associated with an ORFX-associated
CC	pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.	CC	disorder in humans, and in the manufacture of a medicament for treating a
CC	P. acnes is also involved in infections of bone, joints and the central	CC	syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC	nervous system, however it is particularly involved in the inflammatory	CC	sequences can be used in gene therapy. ORFX sequences can be used in the
CC	lesions associated with acne vulgaris. A method for detecting the	CC	treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC	presence or absence of P. acnes in a patient comprises contacting a	CC	psoriasis, benign tumours, keloid, degenerative disorders related to organ
CC	sample with a binding agent that binds to the proteins of the invention	CC	osteoarthritis, neurodegenerative disorders, disorders related to organ
CC	and determining the amount of bound protein in the sample. The	CC	transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC	polypeptides may be used as antigens in the production of antibodies	CC	lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC	specific for P. acnes proteins. These antibodies can be used to	CC	storage disease, various immune deficiencies and disorders, infectious
CC	downregulate expression and activity of P. acnes polypeptides and	CC	diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC	therefore treat P. acnes infections. The antibodies may also be used as	CC	arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC	diagnostic agents for determining P. acnes presence, for example, by	CC	disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC	enzyme linked immunosorbent assay (ELISA).	CC	useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC	Note: The sequence data for this patent did not form part of the printed	CC	bone degenerative disorders, or periodontal disease, and for gut
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 66 AA;		
SQ	Query Match 15.6%; Score 61; DB 22; Length 66;		
	Best Local Similarity 31.8%; Pred. No. 48;		

Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and autoimmune disorders

Disclosure; SEQ ID 2710; 1037pp; English.

The present invention describes substantially purified human proteins (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1 in the specification). ABN15762 to ABN27252 encode the human ORFX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for treating or preventing a pathology associated with an ORFX-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORFX-associated disorder. ORFX polynucleotide sequences can be used in gene therapy. ORFX sequences can be used in the treatment of cancer, hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORFX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut

CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 60 AA;

Query Match 15.3%; Score 60; DB 23; Length 60;

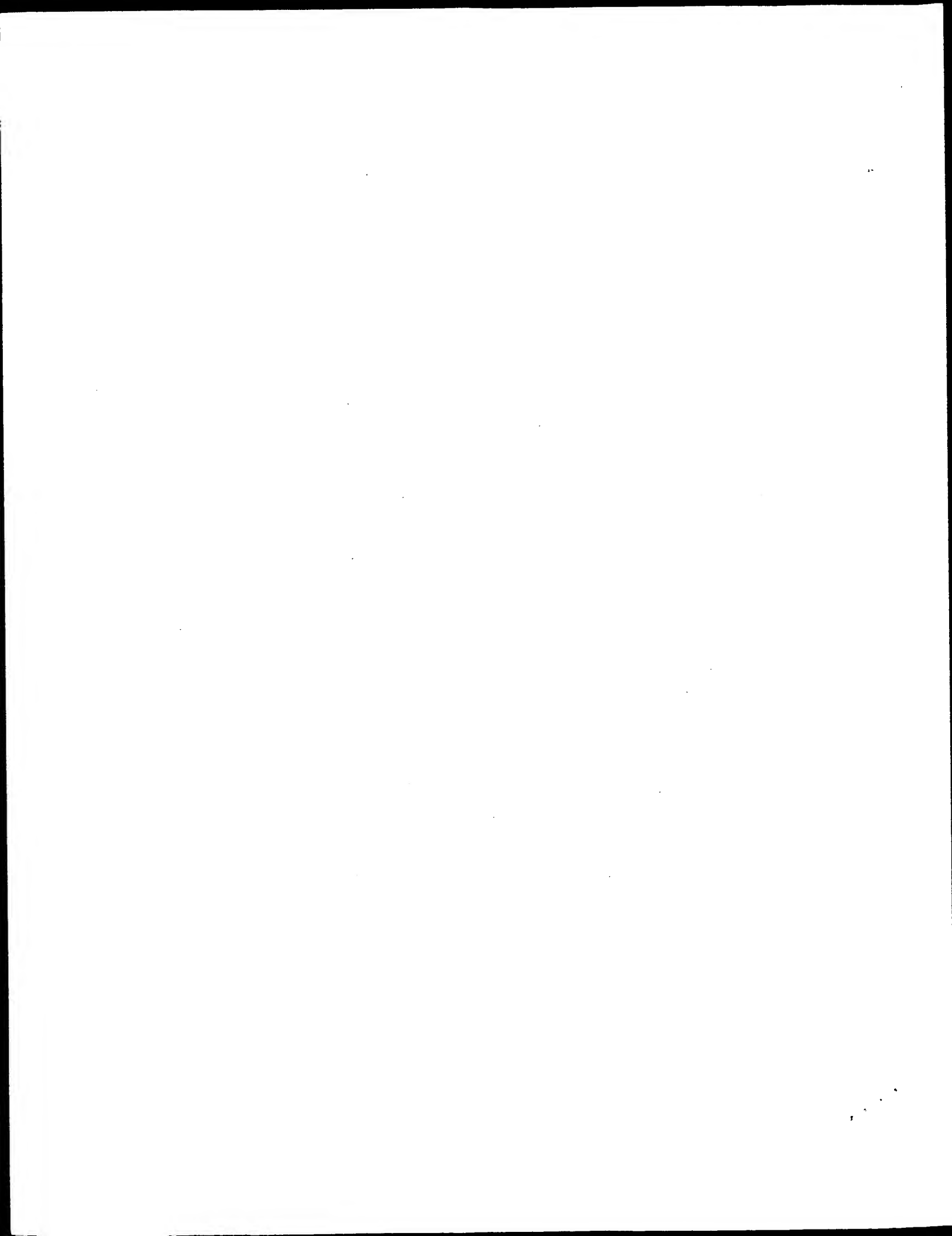
Best Local Similarity 35.5%; Pred. No. 56;

Matches 11; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

Qy 36 PLCIKICAPGVCVCRGLGLRNKKKVCVPRSKC 66

Db 3 PTCACVCRPICACVCVCLLSXVCVCIPIACAC 33

Search completed: February 26, 2003, 15:10:41
Job time : 37 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:12:44 ; Search time 13 Seconds
(without alignments)
194.412 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGLGGRGKCPSEIFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 71782

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	16.1	60	10	US-09-790-264-59
2	62.5	16.0	63	9	US-09-950-933A-76
3	57.5	14.7	61	9	US-09-981-353-115
4	56.5	14.5	61	9	US-09-981-353-120
5	55.5	14.2	38	10	US-09-847-185-47
6	55.5	14.2	45	9	US-10-136-573A-11
7	55.5	14.2	45	9	US-09-877-665-11
8	55.5	14.2	45	9	US-10-215-862-11
9	55.5	14.2	45	10	US-09-817-647-11
10	55.5	14.2	46	9	US-10-201-945-12
11	54.5	13.9	58	10	US-09-865-578-11
12	53	13.6	46	10	US-09-864-761-34995
13	49	12.5	40	10	US-09-917-340-73
14	49	12.5	54	9	US-09-796-692-1773
15	49	12.5	64	9	US-09-796-692-1792
16	49	12.5	64	9	US-09-796-692-2056
17	48.5	12.4	51	10	US-09-864-761-37621
18	48.5	12.4	57	10	US-09-865-578-2
19	48.5	12.4	57	10	US-09-865-578-8

20 48.5 12.4 65 9 US-09-796-692-2465
21 48 12.3 40 10 US-09-917-340-68
22 48 12.3 51 10 US-09-854-864-6
23 48 12.3 64 10 US-09-864-761-47829
24 47.5 12.1 42 10 US-09-865-578-4
25 47 12.0 51 10 US-09-827-948-10
26 47 12.0 53 10 US-09-925-297-901
27 47 12.0 61 10 US-09-829-481-4
28 46.5 11.9 41 10 US-09-864-761-44588
29 46.5 11.9 50 9 US-10-138-158-10
30 46 11.8 37 10 US-09-829-481-11
31 46 11.8 37 10 US-09-917-340-57
32 46 11.8 50 9 US-09-932-172A-2
33 46 11.8 50 9 US-09-955-581-2
34 46 11.8 50 9 US-10-039-119A-2
35 46 11.8 50 9 US-10-138-158-2
36 46 11.8 50 9 US-10-138-158-11
37 46 11.8 50 9 US-10-138-158-12
38 46 11.8 50 10 US-09-299-473A-1
39 45.5 11.6 51 9 US-09-741-106-5
40 45.5 11.6 51 10 US-09-827-948-6
41 45.5 11.6 56 9 US-09-909-567B-52
42 45.5 11.6 63 9 US-09-950-933A-75
43 45.5 11.6 66 10 US-09-965-703-42
44 45 11.5 45 9 US-10-136-573A-13
45 45 11.5 45 9 US-09-877-665-13

ALIGNMENTS

RESULT 1
US-09-790-264-59
; Sequence 59, Application US/09790264
; Patent No. US20020028508A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE,
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-322001
; CURRENT APPLICATION NUMBER: US/09/790,264
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/102,705
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: US 09/363,630
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-790-264-59

Query Match 16.1%; Score 63; DB 10; Length 60;
Best Local Similarity 38.1%; Pred. No. 7.7;
Matches 16; Conservative 2; Mismatches 20; Indels 4; Gaps 2;

QY 10 GKCPSEIFSRCDGRQRCFCPNVVPKPLCIKTCAGCC--VCR 49
||||| | | | | | | | | | | | | | | | | |

Db 17 GKCPNP--PRSIGTCVELSGDSCPNIQKCCSNGCGHVCK 56

RESULT 2

US-09-950-933A-76

; Sequence 76, Application US/09950933A

; Patent No. US20020166141A1

; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.

; APPLICANT: Navarro, Pedro

; TITLE OF INVENTION: Antimicrobial Peptides and Methods of Use

; FILE REFERENCE: 35718/238472

; CURRENT APPLICATION NUMBER: US/09/950,933A

; CURRENT FILING DATE: 2001-09-11

; PRIOR APPLICATION NUMBER: 60/232,569

; PRIOR FILING DATE: 2000-09-13

; NUMBER OF SEQ ID NOS: 99

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 76

; LENGTH: 63

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-09-950-933A-76

Query Match 16.0%; Score 62.5; DB 9; Length 63;

Best Local Similarity 37.8%; Pred. No. 8.9;

Matches 17; Conservative 4; Mismatches 19; Indels 5; Gaps 3;

QY 20 RCDGRCQRCFPCNV-VPKP---LCIKICAPGCVCRGLYLNNKKVC 60

Db 4 QCGGQCTRCSTNTHKHKPCMFQCKCAK-CLCVPPGTGYNKQVC 47

RESULT 3

US-09-981-353-115

; Sequence 115, Application US/09981353

; Patent No. US20020160382A1

; GENERAL INFORMATION:

; APPLICANT: Lasek, Amy W.

; APPLICANT: Jones, David A.

; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER

; FILE REFERENCE: PA-0038 US

; CURRENT APPLICATION NUMBER: US/09/981,353

; CURRENT FILING DATE: 2001-10-11

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PERL Program

; SEQ ID NO 115

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. US20020160382A1 2685676CD1

US-09-981-353-115

Query Match 14.7%; Score 57.5; DB 9; Length 61;

Best Local Similarity 32.6%; Pred. No. 26;

Matches 15; Conservative 6; Mismatches 20; Indels 5; Gaps 4;

QY 4 GGLGGRGKCPSPNEIFSRCDGRCQRCFPCNVVVPKPLCIKICAPGCVCR 49

Db 11 GSCACAGCKCKE--CKCTS-CKKSCCSCPVG-CAK-CAQGCICK 51

RESULT 4

US-09-981-353-120

; Sequence 120, Application US/09981353

; Patent No. US20020160382A1

; GENERAL INFORMATION:

; APPLICANT: Lasek, Amy W.

; APPLICANT: Jones, David A.

; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER

; FILE REFERENCE: PA-0038 US

; CURRENT APPLICATION NUMBER: US/09/981,353

; CURRENT FILING DATE: 2001-10-11

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PERL Program

; SEQ ID NO 120

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. US20020160382A1 2757583CD1

US-09-981-353-120

Query Match 14.5%; Score 56.5; DB 9; Length 61;

Best Local Similarity 37.5%; Pred. No. 33;

Matches 15; Conservative 5; Mismatches 15; Indels 5; Gaps 4;

QY 10 GKCPSPNEIFSRCDGRCQRCFPCNVVVPKPLCIKICAPGCVCR 49

Db 17 GSKCKE--CKCTS-CKKSCCSCPVG-CSK-CAQGCICK 51

RESULT 5

US-09-847-185-47

; Sequence 47, Application US/09847185

; Patent No. US20020076392A1

; GENERAL INFORMATION:

; APPLICANT: Soo Hoo, William

; TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE RESPONSE USING SAME

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CAMPBELL & FLORES, LLP

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92121

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/847,185

; FILING DATE: 01-May-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/201,931

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-IM 2442

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619)535-9001

; TELEFAX: (619)535-8949

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 38 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-847-185-47

Query Match 14.2%; Score 55.5; DB 10; Length 38;

Best Local Similarity 29.8%; Pred. No. 27;

Matches 14; Conservative 0; Mismatches 14; Indels 19; Gaps 2;

QY 2 GFGGLGGRGKCPSPNEIFSRCDGRCQRCFPCNVVVPKPLCIKICAPGCVCR 48

Db 1 GCGGAGGGGCC-----CTAGC-----ACCCACCGGCTC 28

RESULT 6

US-10-136-573A-11
 ; Sequence 11, Application US/10136573A
 ; Patent No. US20020161200A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Mark, Melanie Rose
 ; APPLICANT: Zhang, Dong Xiao
 ; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and
 ; TITLE OF INVENTION: Uses Therefor
 ; FILE REFERENCE: P1084R1C2
 ; CURRENT APPLICATION NUMBER: US 09/480,977
 ; PRIOR APPLICATION NUMBER: US 09/480,977
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: US 08/899,437
 ; PRIOR FILING DATE: 1997-07-24
 ; PRIOR APPLICATION NUMBER: US 60/052,019
 ; PRIOR FILING DATE: 1997-07-09
 ; NUMBER OF SEQ ID NOS: 23
 ; SEQ ID NO 11
 ; LENGTH: 45
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-136-573A-11

Query Match 14.2%; Score 55.5; DB 9; Length 45;
 Best Local Similarity 35.9%; Pred. No. 31;
 Matches 14; Conservative 5; Mismatches 13; Indels 7; Gaps 2;

QY 18 FSRCDGRCQRCFNVVPKPLCIKICA---PGCVCRGLGYL 53
 |||| :
 Db 2 FSRCPKQYKHYC-----IKGRCRFVVAEQTPSCVDEGYI 36

RESULT 7

US-09-877-665-11
 ; Sequence 11, Application US/09877665
 ; Patent No. US20020164680A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
 ; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
 ; TITLE OF INVENTION: Ligands and Uses Therefor

NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/877,665
 FILING DATE: 08-Jun-2001

CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/109,206
 FILING DATE: 30-Jun-1998

ATTORNEY/AGENT INFORMATION:
 NAME: Conley, Deirdre L.
 REGISTRATION NUMBER: 36,487
 REFERENCE/DOCKET NUMBER: P1084R1-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 45 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ; FEATURE:
 ; NAME/KEY: hBTC.efg
 ; LOCATION: 1-45
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION:
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
 US-09-877-665-11

Query Match 14.2%; Score 55.5; DB 9; Length 45;
 Best Local Similarity 35.9%; Pred. No. 31;
 Matches 14; Conservative 5; Mismatches 13; Indels 7; Gaps 2;

QY 18 FSRCDGRCQRCFNVVPKPLCIKICA---PGCVCRGLGYL 53
 |||| :
 Db 2 FSRCPKQYKHYC-----IKGRCRFVVAEQTPSCVDEGYI 36

RESULT 8

US-10-215-862-11
 ; Sequence 11, Application US/10215862
 ; Publication No. US20030036166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Mark, Melanie Rose
 ; APPLICANT: Zhang, Dong Xiao
 ; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and
 ; TITLE OF INVENTION: Uses Therefor

FILE REFERENCE: P1084R1D2C1
 CURRENT APPLICATION NUMBER: US/10/215,862
 CURRENT FILING DATE: 2002-09-24
 PRIOR APPLICATION NUMBER: US 09/126,663
 PRIOR FILING DATE: 1998-07-30
 PRIOR APPLICATION NUMBER: US 08/899,437
 PRIOR FILING DATE: 1997-07-24
 PRIOR APPLICATION NUMBER: US 60/052,019
 PRIOR FILING DATE: 1997-07-09
 NUMBER OF SEQ ID NOS: 23
 ; SEQ ID NO 11
 ; LENGTH: 45
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-215-862-11

Query Match 14.2%; Score 55.5; DB 9; Length 45;
 Best Local Similarity 35.9%; Pred. No. 31;
 Matches 14; Conservative 5; Mismatches 13; Indels 7; Gaps 2;

QY 18 FSRCDGRCQRCFNVVPKPLCIKICA---PGCVCRGLGYL 53
 |||| :
 Db 2 FSRCPKQYKHYC-----IKGRCRFVVAEQTPSCVDEGYI 36

RESULT 9

US-09-817-647-11
 ; Sequence 11, Application US/09817647
 ; Patent No. US20020082229A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
 ; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
 ; TITLE OF INVENTION: Ligands and Uses Therefor

NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA

```

; APPLICATION NUMBER: US 08/208,008
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF11001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-201-945-12

Query Match 14.2%; Score 55.5; DB 9; Length 46;
Best Local Similarity 35.9%; Pred. No. 32;
Matches 14; Conservative 5; Mismatches 13; Indels 7; Gaps 2;

QY 18 FSRGDCRCQFCFNNVPKPLCIKICA---PGCVCRGLYL 53
|||| : : | | : | | | | | | |
Db 3 FSRCPKYKHYC-----IKGRCFVVAEQTPSCVCDGYI 37

RESULT 11
US-09-865-578-11
; Sequence 11, Application US/09865578
; Patent No. US20010034433A1
; GENERAL INFORMATION:
; APPLICANT: E. MARTIN, SPENCER
; TITLE OF INVENTION: HUMAN SOMATOMEDIAN CARRIER PROTEIN SUBUNITS
; TITLE OF INVENTION: AND PROCESS FOR PRODUCING THEM; RECOMBINANT DNA MOLECULES
; TITLE OF INVENTION: HOSTS, PROCESSES AND HUMAN SOMATOMEDIAN CARRIER PROTEIN-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 057491/0680
; CURRENT APPLICATION NUMBER: US/09/865,578
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/397,192
; PRIOR FILING DATE: 1999-09-16
; PRIOR APPLICATION NUMBER: 09/162,118
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: 08/923,860
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 08/706,755
; PRIOR FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/437,407
; PRIOR FILING DATE: 1995-05-12
; PRIOR APPLICATION NUMBER: 08/320,123
; PRIOR FILING DATE: 1994-10-07
; PRIOR APPLICATION NUMBER: 08/043,039
; PRIOR FILING DATE: 1993-04-05
; PRIOR APPLICATION NUMBER: 07/763,481
; PRIOR FILING DATE: 1991-09-20
; PRIOR APPLICATION NUMBER: 07/290,250
; PRIOR FILING DATE: 1988-12-22
; PRIOR APPLICATION NUMBER: 07/170,022
; PRIOR FILING DATE: 1988-03-31
; PRIOR APPLICATION NUMBER: 07/034,885
; PRIOR FILING DATE: 1987-04-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 58
; TYPE: PPT
; ORGANISM: Homo Sapiens
US-09-865-578-11

Query Match 13.9%; Score 54.5; DB 10; Length 58;
Best Local Similarity 34.6%; Pred. No. 49;

```


Matches 18; Conservative 2; Mismatches 15; Indels 17; Gaps 4;

QY 5 GLGGRGKCPNSIFSRCDRCQRCFCNVVPKPLCIKICA-----PGCVCRLG 51
Db 7 GLGPPVRC-----EPCDARALAAQC---APPP---AVCAELVREPGCCCLG 46

RESULT 12

US-09-864-761-34995
; Sequence 34995, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34995
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB020862.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.9
US-09-864-761-34995

Query Match 13.6%; Score 53; DB 10; Length 46;
Best Local Similarity 48.1%; Pred. No. 55;
Matches 13; Conservative 1; Mismatches 5; Indels 8; Gaps 1;

QY 1 GGFGGLGGRGKCPNSIFSRCDRCQRCOR 27
Db 10 GGAGGGRGGRKTR-----DGRIR 28

RESULT 13

US-09-917-340-73
; Sequence 73, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McAnulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; PRIOR FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-917-340-73

Query Match 12.5%; Score 49; DB 10; Length 40;
Best Local Similarity 48.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 42 CAPGCVCR--LGYLRNKKKVCVPRS 64
Db 16 CAAHCLARGNRGGYCNKSKVCVCRN 40

RESULT 14

US-09-796-692-1773
; Sequence 1773, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:10:48 : Search time 352 Seconds
(without alignments)
122.719 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGRGKCPSEIFSR.....CRLGYLRNKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 2759142

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	391	100.0	67	19	US-09-506-978-1
2	391	100.0	67	25	US-10-174-151-1
3	391	100.0	67	26	US-10-204-145-1
4	113.5	29.0	62	27	US-60-142-896-1491
5	113.5	29.0	62	27	US-60-145-138-701
6	88.5	22.6	62	21	US-09-791-537-9664

7	87	22.3	39	27	US-60-160-203-3377	Sequence 3377, Ap
8	87	22.3	39	27	US-60-169-840-5102	Sequence 5102, Ap
9	86.5	22.1	54	27	US-60-162-247-3339	Sequence 3339, Ap
10	84.5	21.6	53	27	US-60-163-123-1354	Sequence 1354, Ap
11	84.5	21.6	53	27	US-60-163-123-1440	Sequence 1440, Ap
12	84	21.5	60	27	US-60-160-203-4074	Sequence 4074, Ap
13	84	21.5	60	27	US-60-163-123-1552	Sequence 1552, Ap
14	84	21.5	60	27	US-60-163-123-1677	Sequence 1677, Ap
15	82.5	21.1	63	27	US-60-169-840-5359	Sequence 5359, Ap
16	82	21.0	61	16	US-09-249-473-61	Sequence 61, Appl
17	82	21.0	61	18	US-09-498-272-61	Sequence 61, Appl
18	82	21.0	61	18	US-09-498-556-61	Sequence 61, Appl
19	80.5	20.6	61	21	US-09-791-537-13914	Sequence 13914, A
20	80.5	20.6	63	21	US-09-791-537-86288	Sequence 86288, A
21	78.5	20.1	65	21	US-09-791-537-86289	Sequence 86289, A
22	77.5	19.8	63	21	US-09-791-537-4553	Sequence 4553, Ap
23	76	19.4	52	27	US-60-162-243-691	Sequence 691, App
24	75.5	19.3	64	27	US-60-196-174-1132	Sequence 1132, Ap
25	73	18.7	43	27	US-60-164-762-848	Sequence 848, App
26	73	18.7	60	21	US-09-791-537-125018	Sequence 125018, A
27	72	18.4	43	27	US-60-164-762-1114	Sequence 1114, Ap
28	71.5	18.3	50	27	US-60-160-203-3591	Sequence 3591, Ap
29	71.5	18.3	53	27	US-60-160-203-3421	Sequence 3421, Ap
30	70.5	18.0	62	21	US-09-791-537-103884	Sequence 103884, A
31	70	17.9	53	27	US-60-170-373-3452	Sequence 3452, Ap
32	69.5	17.8	51	27	US-60-163-123-1638	Sequence 1638, Ap
33	69.5	17.8	56	27	US-60-188-162-2853	Sequence 2853, Ap
34	69	17.6	55	9	US-08-548-186-196	Sequence 196, App
35	69	17.6	60	21	US-09-791-537-137772	Sequence 137772, A
36	69	17.6	60	21	US-09-791-537-137774	Sequence 137774, A
37	68.5	17.5	52	21	US-09-791-537-90125	Sequence 90125, A
38	68	17.4	58	16	US-09-249-473-60	Sequence 60, Appl
39	68	17.4	58	18	US-09-498-272-60	Sequence 60, Appl
40	68	17.4	58	18	US-09-498-556-60	Sequence 60, Appl
41	68	17.4	60	21	US-09-791-537-31406	Sequence 31406, A
42	68	17.4	60	21	US-09-791-537-44038	Sequence 44038, A
43	67	17.1	62	21	US-09-791-537-51265	Sequence 51265, A
44	67	17.1	62	27	US-60-389-987-1874	Sequence 1874, Ap
45	67	17.1	62	27	US-60-412-418-1874	Sequence 1874, Ap

ALIGNMENTS

RESULT 1

```

US-09-506-978-1
; Sequence 1, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/09/506,978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-09-506-978-1

```

Query Match 100.0%; Score 391; DB 19; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.le-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGFGLGGRGKCPSEIFSRCDGRQRCFCNPVVPKPLCIKICAPGCVRLGYLRNKKVC 60
Db 1 GGFGLGGRGKCPSEIFSRCDGRQRCFCNPVVPKPLCIKICAPGCVRLGYLRNKKVC 60
QY 61 VPRSKG 67
Db 61 VPRSKG 67

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:11:03 ; Search time 24 Seconds
(without alignments)
255.149 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGLGGRGKCPNSIEFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 231224

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	30.4	67	6	US-10-038-854-185
2	63	16.1	60	6	US-10-289-353-59
3	62.5	16.0	62	6	US-10-285-876-5
4	61	15.6	66	1	PCT-US02-32727-6458
5	61	15.6	66	5	US-09-978-825-6458
6	61	15.6	66	6	US-10-057-498-6458
7	58.5	15.0	61	6	US-10-170-385-239
8	58	14.8	56	1	PCT-US02-32727-9993
9	58	14.8	56	5	US-09-978-825-9993
10	58	14.7	56	6	US-10-057-498-9993
11	57.5	14.7	61	6	US-10-170-385-241
12	57.5	14.7	61	7	US-60-440-068-445
13	57	14.6	67	5	US-09-724-676-81522
14	57	14.6	67	5	US-09-724-676A-81522
15	56.5	14.5	61	6	US-10-170-385-265
16	56.5	14.5	61	7	US-60-423-586-136
17	56.5	14.5	61	7	US-60-427-194-136
18	55.5	14.2	46	5	US-09-857-815A-4
19	55.5	14.2	47	5	US-09-857-815A-3
20	55.5	14.2	47	5	US-09-857-815A-12
21	55.5	14.2	48	5	US-09-857-815A-11
22	55.5	14.2	48	5	US-09-857-815A-14
23	55.5	14.2	49	5	US-09-857-815A-10
24	55.5	14.2	49	5	US-09-857-815A-13
25	55.5	14.2	58	6	US-10-231-778-229
26	55	14.1	66	1	PCT-US02-32727-16329

ALIGNMENTS

RESULT 1

US-10-038-854-185

; Sequence 185, Application US/10038854

; GENERAL INFORMATION: Spytek, Kimberly A

; APPLICANT: Li, Li

; APPLICANT: Wolenc, Adam R

; APPLICANT: Vernet, Corine

; APPLICANT: Eisen, Andrew J

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Shimkets, Richard A

; APPLICANT: Tchernev, Velizar

; APPLICANT: Gorman, Linda

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Gangolli, Esha A

; APPLICANT: Guo, Xiaojia S

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J

; APPLICANT: Boldog, Ferenc

; APPLICANT: Burgess, Catherine E

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Ellerman, Karen

; APPLICANT: Gunther, Erik

; APPLICANT: Smithson, Glenda

; APPLICANT: Millet, Isabelle

; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-230

; CURRENT APPLICATION NUMBER: US/10/038,854

; CURRENT FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928

; PRIOR FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: 60/259,415

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: 60/259,785

; PRIOR FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 60/269,814

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,863

; PRIOR FILING DATE: 2001-03-29

Sequence 16329, A
Sequence 16329, A
Sequence 6751, Ap
Sequence 6751, Ap
Sequence 44, Appl
Sequence 12069, A
Sequence 243, App
Sequence 81494, A
Sequence 81494, A
Sequence 20050, A
Sequence 20050, A
Sequence 20050, A
Sequence 83, Appl
Sequence 3458, Ap
Sequence 233, App
Sequence 47, Appl
Sequence 4852, Ap
Sequence 4852, Ap

27 55 14.1 66 5 US-09-978-825-16329
28 55 14.1 66 6 US-10-057-498-16329
29 54.5 13.9 60 1 PCT-US02-32727-6751
30 54.5 13.9 60 5 US-09-978-825-6751
31 54.5 13.9 60 6 US-10-057-498-6751
32 54 13.8 53 5 US-09-857-815A-44
33 53 13.6 46 6 US-10-203-138A-12069
34 53 13.6 61 6 US-10-170-385-243
35 53 13.6 67 5 US-09-724-676-81494
36 53 13.6 67 5 US-09-724-676A-81494
37 52.5 13.4 57 1 PCT-US02-32727-20050
38 52.5 13.4 57 5 US-09-978-825-20050
39 52.5 13.4 57 6 US-10-057-498-20050
40 51 13.0 36 6 US-10-004-378A-83
41 51 13.0 61 6 US-10-218-140-3458
42 50.5 12.9 57 6 US-10-231-778-233
43 50 12.8 36 1 PCT-US02-18256-47
44 50 12.8 58 1 PCT-US02-32727-4852
45 50 12.8 58 5 US-09-978-825-4852

; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 6458
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-6458

Query Match 15.6%; Score 61; DB 1; Length 66;
Best Local Similarity 31.8%; Pred. No. 17;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;

QY 2 GFGGLGGRGKC---PSNEIFSRCDGRCPNVVVPKPLCIKICAPGCVCRGLYLNRNKKK 58
Db 7 GRGELISAGKCGFRPSDQ--RRCSAQCGGRLPGSRPVSRCVR-----CIC--CSLRASPL 57
QY 59 VCVPRS 64
Db 58 PCQPPS 63

RESULT 5

US-09-978-825-6458
; Sequence 6458, Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqian
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: US/09/978,825
; CURRENT FILING DATE: 2003-01-29
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 6458
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
US-09-978-825-6458

Query Match 15.6%; Score 61; DB 5; Length 66;
Best Local Similarity 31.8%; Pred. No. 17;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;

QY 2 GFGGLGGRGKC---PSNEIFSRCDGRCPNVVVPKPLCIKICAPGCVCRGLYLNRNKKK 58
Db 7 GRGELISAGKCGFRPSDQ--RRCSAQCGGRLPGSRPVSRCVR-----CIC--CSLRASPL 57
QY 59 VCVPRS 64
Db 58 PCQPPS 63

RESULT 6

US-10-057-498-6458
; Sequence 6458, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 6458
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
US-10-057-498-6458

Query Match 15.6%; Score 61; DB 6; Length 66;
Best Local Similarity 31.8%; Pred. No. 17;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;
QY 2 GFGGLGGRGKC---PSNEIFSRCDGRCPNVVVPKPLCIKICAPGCVCRGLYLNRNKKK 58
Db 7 GRGELISAGKCGFRPSDQ--RRCSAQCGGRLPGSRPVSRCVR-----CIC--CSLRASPL 57
QY 59 VCVPRS 64
Db 58 PCQPPS 63

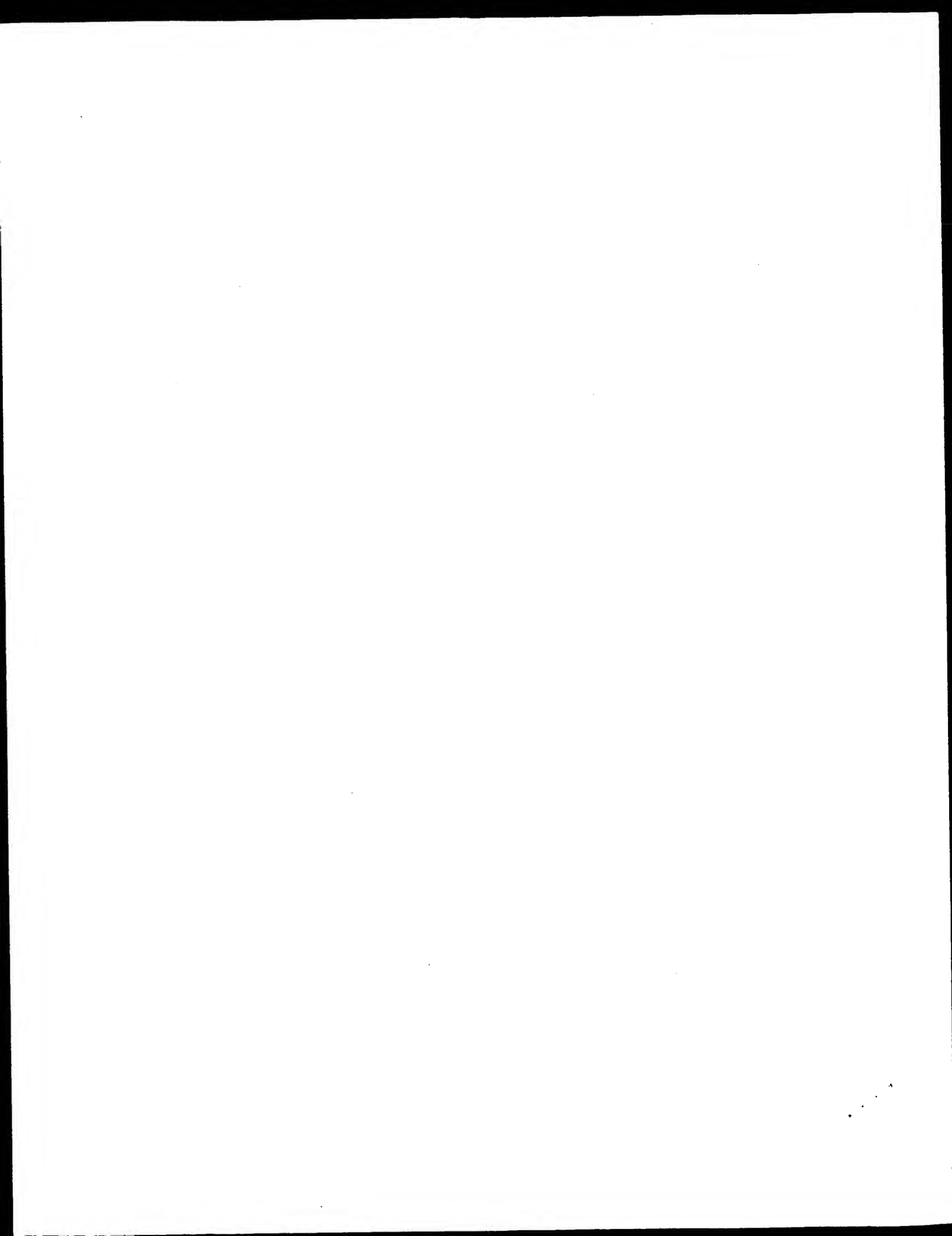
RESULT 7

US-10-170-385-239
; Sequence 239, Application US/10170385
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 239
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-239

Query Match 15.0%; Score 58.5; DB 6; Length 61;
Best Local Similarity 32.7%; Pred. No. 29;
Matches 16; Conservative 6; Mismatches 20; Indels 7; Gaps 4;
QY 1 GFGGLGGRGKCPSNEIFSRCDGRCPNVVVPKPLCIKICAPGCVCR 49
Db 10 GGCACAGSCKCKK---CKCTS-CKKSCCSCCPPLG-CAK-CAQGCICK 51

RESULT 8

PCT-US02-32727-9993
; Sequence 9993, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:10:07 ; Search time 14 Seconds
(without alignments)
140.810 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGGGGLGGGKCPSEIFSR.....CRLGYLRNKKKVCVPSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 188167

Minimum DB seq length: 0
Maximum DB seq length: 67

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	82	21.0	61	2	US-08-465-380-61
2	82	21.0	61	2	US-08-486-397-61
3	82	21.0	61	2	US-08-486-399-61
4	82	21.0	61	2	US-08-461-965-61
5	82	21.0	61	2	US-08-634-641-61
6	82	21.0	61	3	US-09-249-471-61
7	82	21.0	61	3	US-09-249-472-61
8	82	21.0	61	3	US-09-249-451-61
9	82	21.0	61	3	US-08-809-455-61
10	82	21.0	61	3	US-09-249-461-61
11	82	21.0	61	3	US-09-249-448-61
12	68	17.4	47	2	US-08-637-759B-400
13	68	17.4	47	3	US-08-871-355A-400
14	68	17.4	47	4	US-09-201-945-400
15	68	17.4	58	2	US-08-465-380-60
16	68	17.4	58	2	US-08-486-397-60
17	68	17.4	58	2	US-08-486-399-60
18	68	17.4	58	2	US-08-461-965-60
19	68	17.4	58	2	US-08-634-641-60
20	68	17.4	58	3	US-09-249-471-60
21	68	17.4	58	3	US-09-249-472-60
22	68	17.4	58	3	US-09-249-451-60
23	68	17.4	58	3	US-08-809-455-60
24	68	17.4	58	3	US-09-249-461-60
25	68	17.4	58	3	US-09-249-448-60
26	62.5	16.0	62	4	US-07-780-717C-5
27	59.5	15.2	61	2	US-08-785-530-3

28 59.5 15.2 61 2 US-09-123-850-3
29 59 15.1 61 2 US-08-785-530-1
30 59 15.1 61 2 US-09-123-850-1
31 57.5 14.7 61 2 US-08-785-530-4
32 57.5 14.7 61 2 US-09-123-850-4
33 57 14.6 10 4 US-09-394-630-20
34 55.5 14.2 38 2 US-08-902-516-47
35 55.5 14.2 45 3 US-08-899-437-11
36 55.5 14.2 45 4 US-09-126-121-11
37 55.5 14.2 46 4 US-08-915-096A-12
38 54.5 13.9 61 2 US-08-785-530-6
39 54.5 13.9 61 2 US-09-123-850-6
40 54 13.8 61 2 US-08-785-530-5
41 54 13.8 61 2 US-09-123-850-5
42 54 13.8 63 2 US-08-369-829A-18
43 52.5 13.4 44 6 5177197-48
44 52 13.3 42 2 US-08-761-248B-15
45 51.5 13.2 48 2 US-08-465-794-3

ALIGNMENTS

RESULT 1
US-08-465-380-61
; Sequence 61, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb.
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Necator americanus

US-08-465-380-61

Query Match 21.0%; Score 82; DB 2; Length 61;
Best Local Similarity 30.9%; Pred. No. 0.051;
Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;

Qy 12 CPSNEIFSRCDGRCQFCPNVVPKPLCIKICAPGV-----CRIGYLNRNKKKVCV 61
Db 4 CPANEEWRECGTPECPKCNQMP-----DICTMNCIVDVCCKEGYKRHETKGCL 53

RESULT 2

US-08-486-397-61

Sequence 61, Application US/08486397

Patent No. 5866342

GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

NUMBER OF SEQUENCES: 357

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,397

FILING DATE: June 5, 1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 213/270

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

FILING DATE: June 5, 1995

SEQUENCE CHARACTERISTICS:

LENGTH: 61 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Necator americanus

US-08-486-399-61

Query Match 21.0%; Score 82; DB 2; Length 61;
Best Local Similarity 30.9%; Pred. No. 0.051;
Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;

Qy 12 CPSNEIFSRCDGRCQFCPNVVPKPLCIKICAPGV-----CRIGYLNRNKKKVCV 61
Db 4 CPANEEWRECGTPECPKCNQMP-----DICTMNCIVDVCCKEGYKRHETKGCL 53

RESULT 4

US-08-461-965-61

Sequence 61, Application US/08461965

Patent No. 5872098

GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,399

FILING DATE: June 5, 1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 213/270

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

FILING DATE: June 5, 1995

SEQUENCE CHARACTERISTICS:

LENGTH: 61 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Necator americanus

US-08-486-397-61

US-08-465-380-61

Query Match 21.0%; Score 82; DB 2; Length 61;
Best Local Similarity 30.9%; Pred. No. 0.051;
Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;

Qy 12 CPSNEIFSRCDGRCQFCPNVVPKPLCIKICAPGV-----CRIGYLNRNKKKVCV 61
Db 4 CPANEEWRECGTPECPKCNQMP-----DICTMNCIVDVCCKEGYKRHETKGCL 53

RESULT 2

US-08-486-397-61

Sequence 61, Application US/08486397

Patent No. 5866342

GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

NUMBER OF SEQUENCES: 357

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,397

FILING DATE: June 5, 1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158

REFERENCE/DOCKET NUMBER: 213/269

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

FILING DATE: June 5, 1995

SEQUENCE CHARACTERISTICS:

LENGTH: 61 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Necator americanus

US-08-486-397-61

Query Match 21.0%; Score 82; DB 2; Length 61;
Best Local Similarity 30.9%; Pred. No. 0.051;
Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;

Qy 12 CPSNEIFSRCDGRCQFCPNVVPKPLCIKICAPGV-----CRIGYLNRNKKKVCV 61
Db 4 CPANEEWRECGTPECPKCNQMP-----DICTMNCIVDVCCKEGYKRHETKGCL 53

RESULT 3


```

RESULT 7
US-09-249-472-61
; Sequence 61, Application US/09249472
; Patent No. 6046318
; GENERAL INVENTION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon

```

RESULT 8

APPLICANT: Stanssens, Patrick Eric Hugo
 APPLICANT: Messens, Joris Hilda Llieven
 APPLICANT: Lauwereys, Marc Josef
 APPLICANT: Laroche, Yves Rene
 APPLICANT: Jespers, Laurent Stephane
 APPLICANT: Gansemans, Yannick Georges Jozef
 APPLICANT: Moyle, Matthew
 APPLICANT: Berdum, Peter W.
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
 TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
 TITLE OF INVENTION: PROTEIN
 NUMBER OF SEQUENCES: 356
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/249,448
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/809,455
 FILING DATE: April 17, 1997
 APPLICATION NUMBER: PCT/US95/13231
 FILING DATE: October 17, 1995
 APPLICATION NUMBER: 08/486,399
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/486,397
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/465,380
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/461,965
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 216/270
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 61:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 61 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Necator americanus
 IS-09-249-448-61

Query Match	21.0%;	Score 82;	DB 3;	Length 61;
Best Local Similarity	30.9%;	Pred. No. 0.051;		
Matches	17:	Conservative	9:	Mismatches 19; Indels 10; Gaps 2;

[illegible]

RESULT 12
US-08-637-759B-400

; Sequence 400, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-400

Query Match 17.4%; Score 68; DB 2; Length 47;
Best Local Similarity 55.6%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCORFC 29
||: ||||| |
Db 30 CPAGKPLSRCDGRCDEIC 47

RESULT 13
US-08-871-355A-400
; Sequence 400, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-871-355A-400

Query Match 17.4%; Score 68; DB 3; Length 47;
Best Local Similarity 55.6%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCORFC 29
||: ||||| |
Db 30 CPAGKPLSRCDGRCDEIC 47

RESULT 14
US-09-201-945-400
; Sequence 400, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:

;
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-400

Query Match

Best Local Similarity 17.4%; Score 68; DB 4; Length 47;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCQRCFC 29
||:..||||||| |
Db 30 CPAGPLSRCDGRCDEIC 47

RESULT 15

US-08-465-380-60
; Sequence 60, Application US/08465380
; Patent No. 3863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071

COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 60:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Heligmosomoides polygyrus
US-08-465-380-60

Query Match

Best Local Similarity 17.4%; Score 68; DB 2; Length 58;
Matches 14; Conservative 7; Mismatches 34; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCQRCFCNNVVKPLCLIKICAPGCVCRGLYLNNKKKVCVPRSKC 66
||:..||||||| |
Db 3 CGPNEEYTECGTCEPKCNEPMDICTLNCIVNVCQCKPCKRGPCKGCVAPGPGC 57

Search completed: February 26, 2003, 15:13:22
Job time : 15 secs